UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS FÍSICAS

TESIS DOCTORAL

New soft-computing algorithms in atmospheric physics

Nuevos algoritmos de soft-computing en física atmosférica

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Sancho Salcedo Sanz

Director

Ricardo Francisco García Herrera

Madrid
Ed. electrónica 2019

© **Sancho Salcedo Sanz, 2019**

Programa de Doctorado en Física

NEW SOFT-COMPUTING ALGORITHMS IN ATMOSPHERIC PHYSICS

Tesis Doctoral presentada por

SANCHO SALCEDO SANZ

2019

UNIVERSIDAD COMPLUTENSE MADRID

Programa de Doctorado en Física

NEW SOFT-COMPUTING ALGORITHMS IN ATMOSPHERIC PHYSICS

Tesis Doctoral presentada por

SANCHO SALCEDO SANZ

Director:
PROF. DR. RICARDO FRANCISCO GARCÍA HERRERA

Madrid, 2019

UNIVERSIDAD COMPLUTENSE MADRID

Universidad Complutense de Madrid
Dpto. de Física de la Tierra y Astrofísica
Plaza de Ciencias, 1
28040 Madrid
Telf: +34 91 394 44 90
Fax: +34 91 394 46 35

Dr. D. Ricardo Francisco García Herrera, Catedrático de Universidad del Departamento de Física de la Tierra y Astrofísica de la Universidad Complutense de Madrid,

CERTIFICA

Que la Tesis Doctoral titulada "**New Soft-Computing Algorithms in Atmospheric Physics**", presentada por D. Sancho Salcedo Sanz y realizada bajo mi dirección, reúne méritos suficientes para optar al grado de Doctor, por lo que puede procederse a su depósito y defensa.

Madrid, 14 de enero de 2019.

Fdo.: Dr. D. Ricardo Francisco García Herrera

Esta Tesis Doctoral está dedicada al Profesor Emiliano Hernández Martín,
Maestro de maestros.

Abstract

This Ph.D. Thesis elaborates and analyzes several hybrid Soft-Computing algorithms for optimization and prediction problems in Atmospheric Physics. The core of the Thesis is a recently developed optimization meta-heuristic, the Coral Reefs Optimization Algorithm (CRO), an evolutionary-based approach which considers a population of possible solutions to a given optimization problem. It simulates different procedures mimicking real processes occurring in coral reefs in order to evolve the population towards good solutions for the problem. Alternative modifications of this algorithm lead to powerful co-evolution meta-heuristics, such as the CRO-SL, in which *Substrates* implementing different search procedures are included. Another modification of the algorithm leads to the CRO-SP, which considers *Species* in the evolution of the population, and it is able to deal with different encodings within a single population. These approaches are hybridized with other Machine Learning and traditional algorithms such as neural networks or the Analogue Method (AM), to come up with powerful hybrid approaches able to solve hard problems in Atmospheric Physics.

Different applications are tackled with these approaches:

- Feature selection problems for machine learning prediction algorithms. We deal with problems of selecting the best set of features which mostly improve the performance of a given regressor, usually a fast-training neural network such as Extreme Learning Machines (ELMs). We evaluate different alternatives for this, such as hybrid CRO-ELM, CRO-SL-ELM or CRO-SP-ELM, in comparison with other possibilities. With these algorithms we tackle two problems of wind speed prediction and a problem of global solar radiation prediction. In both applications the hybrid algorithms have obtained results which improve the performance of alternative algorithms, and of those systems without feature selection process.

- Representative Selection of measuring points for optimal field reconstruction of atmospheric variables. In this case, we state this problem as an integer-encoding optimization task, in which the objective function is a measure of the error field reconstruction given by the AM. We evaluate the performance of the CRO-SL in this problem, with two specific applications: representative selection of the best points to reconstruct air temperature and wind speed fields in Europe. In the case of temperature, we show that the method is able

to successfully deal with gridded and un-gridded datasets. The reconstructions obtained from the selected points have a high quality, and the points selected are climatologically significant. In the case of the wind speed field, we only deal with gridded data from ERA-Interim Reanalysis. The results obtained have shown that the best representative points are mainly located over the North Atlantic Ocean.

These methodologies are fully detailed and analyzed in the context of previous works in the two first chapters of the Thesis, whereas there are four chapters devoted to explore the performance of the hybrid approach in the different applications considered. The work is closed by a final Conclusions and Remarks section, where future lines of research are also discussed.

Resumen en Castellano

En esta Tesis Doctoral se elaboran y analizan en detalle diferentes algoritmos híbridos de *Soft-Computing* para problemas de optimización y predicción en Física de la Atmósfera. El núcleo central de la Tesis es un algoritmo meta-heurístico de optimización recientemente desarrollado, conocido como *Coral Reefs Optimization algorithm* (CRO). Este algoritmo pertenece a la familia de la Computación Evolutiva, de forma que considera una población de soluciones a un problema concreto, y simula los diferentes procesos que ocurren en un arrecife de coral para evolucionar dicha población hacia la solución óptima del problema. Recientemente se han propuesto diferentes versiones del algoritmo CRO básico para obtener mecanismos potentes de optimización co-evolutiva. Una de estas modificaciones es el CRO-SL, en la que se definen un conjunto de *Sustratos* en el algoritmo, de manera que cada sustrato simula un mecanismo de evolución diferente, que son aplicados a la vez en una única población. Otra modificación ha dado lugar al conocido como CRO-SP, un algoritmo donde se definen diferentes *Especies*, capaz de manejar varias codificaciones para un mismo problema a la vez. Estas versiones del CRO han sido hibridadas con varias técnicas de Aprendizaje Máquina, tales como varios tipos de redes neuronales de entrenamiento rápido, sistemas de aprendizaje tales como Máquinas de Vectores Soporte, o sistemas de predicción vinculados totalmente al área de la Física Atmosférica, tales como el Método de los Análogos (AM). Los algoritmos híbridos obtenidos son muy robustos y capaces de obtener excelentes soluciones en diferentes problemas donde han sido probados. Específicamente:

- Problemas de Selección de Características para algoritmos de predicción basados en Aprendizaje Máquina. Este tipo de problemas tiene como objetivo el obtener un subconjunto de las mejores características (variables de entrada) en un sistema de predicción, usualmente dominado por un algoritmo de aprendizaje de tipo neuronal o similar. La idea es obtener el mejor conjunto de características de forma que el rendimiento del sistema de predicción se maximice. En el caso particular discutido en esta Tesis, la predicción se lleva a cabo a partir de redes de entrenamiento rápido, fundamentalmente Máquinas de Aprendizaje Extremo (ELMs). En la Tesis se evalúan diferentes alternativas para estos problemas de Selección de Características, tales como modelos híbridos CRO-ELM, CRO-SL-ELM o CRO-SP-ELM, en contraposición con otro tipo de predictores alternativos y sistemas sin estructuras de Selección de Características. Los algoritmos híbridos se evalúan en dos

problemas de Selección de Características en predicción de viento y en un problema de predicción de radiación solar global. En todas las aplicaciones los algoritmos híbridos han obtenido excelentes resultados, mejorando el rendimiento de las técnicas alternativas con las que han sido comparados.

- Problemas de Selección de puntos Representativos para la reconstrucción óptima de campos de variables atmosféricas. Este caso puede ser tratado como un problema de optimización con codificación entera, en el que la función objetivo viene dada por la salida del Método de los Análogos, que se usa para realizar la reconstrucción del campo en función de la información de los puntos representativos seleccionados. Se evalúa en detalle el rendimiento del algoritmo CRO-SL para este problema, donde un conjunto de operadores que funcionan bien en problemas con codificación entera ha sido seleccionado. Se han abordado dos problemas diferentes, uno de ellos sobre obtención de los puntos representativos para reconstrucción del campo de temperatura y un segundo basado en campos de velocidad de viento. En el primero de ellos se evalúa el algoritmo, mostrándose que puede funcionar con alto rendimiento en datos sobre grids regulares o irregulares. Se discuten asimismo ciertas peculiaridades del método, como su capacidad para encontrar los puntos con menos información para la reconstrucción, simplemente maximizando la salida de la función objetivo, o la coherencia climatológica de los resultados obtenidos. En el caso del campo de velocidad de viento, se estiman los mejores puntos para la reconstrucción del campo, lo que puede tener aplicaciones en energía eólica.

Las diferentes metodologías híbridas presentadas en la Tesis son descritas con detalle y analizadas en profundidad en los dos primeros capítulos, donde también se ofrece un estudio sobre trabajos previos en los ámbitos de algoritmia y aplicación discutidos en la Tesis. Los siguientes capítulos están centrados en la aplicación de los algoritmos, tanto en problemas de Selección de Características como en el problema de Selección de puntos Representativos en campos de variables atmosféricas. La Tesis se cierra con un capítulo final de conclusiones y líneas futuras de trabajo, donde es establece la posible continuidad de esta investigación.

Acknowledgements

- The research of this Ph.D. Thesis has been partially supported by projects TIN2014-54583-C2-2-R and TIN2017-85887-C2-2-P of the *Spanish Ministerial Commission of Science and Technology (MICYT)*, and by *Comunidad de Madrid*, under project number S2013/ICE-2933.

- We acknowledge the use of ECA&D and HISTALP projects' data in Chapter 5. Data and metadata are available at http://www.ecad.eu and http://www.zamg.ac.at/histalp, respectively. ERA-Interim reanalysis data were used in Chapters 5 and 6. The basic reference to ERA-Interim reanalysis is [Dee2011]. Wind speed data used in Chapter 3 were provided by Iberdrola. Solar radiation data used in Chapter 4 were obtained from AEMET, whereas the WRF variables were obtained by the Remote Sensing Laboratory (LATUV, Universidad de Valladolid).

- The rights of the images that appear in the cover of this Thesis belong to the following film distributors/authors: Warner Bros. Pictures (A.I.), Lightning Entertainment (The Reef), Carl-A. Fechner (Die 4. Revolution), Loews Cineplex Entertainment (Gone with the Wind), and Fox Searchlight Pictures (The Sunshine).

Agradecimientos

Hace 16 años estaba a unos pocos kilómetros de aquí escribiendo algo parecido a esto. Todo el mundo sabe que 16 años no son nada: tan cierto como que vivimos en *Matrix*... y como que mis nuevos doctorandos son muuucho peores que los de antes, por supuesto. La verdad es que, para no ser nada, este tiempo ha dado bastante de sí. Lo suficiente como para que un buen día me plantera dejar de ser el *prota* de la canción de X-ambassadors, e intentara agenciarme una nueva muceta de color azul. Más que nada porque la color caldero con la que voy a las aperturas de curso en Alcalá es fea. Pero fea con ganas. Además el gorrito con dos colores lo va a petar, claramente. Bueno, también está el hecho de que me apetecía mucho volver a los orígenes en la investigación, que los problemas en Física de la Atmósfera son verderamente retadores y, finalmente, que tenía muchas ganas de poder mostrar que los algoritmos de Soft-Computing tienen bastante que decir en este campo. Pero no hay que desenfocarse: estos son detalles menores comparados con los que he mencionado anteriormente.

Como no podía ser de otra manera, ya puestos a meterme en este jardín, al menos hacerlo al lado de uno de los mayores expertos en Física de la Atmósfera: Mi director de Tesis, el Dr. Ricardo García Herrera. Gracias, Ricardo, por tus consejos, tus sabias interpretaciones, tu increible intuición y tu guía en el proceso: no sólo durante este tiempo reciente, sino desde que nos conocimos cuando yo era estudiante de Licenciatura. La verdad es que en este nuevo tiempo de Tesis me he vuelto a sentir un poco como mis queridos *mindundis* Laura y Carlos, (sin tranchetes, no os paséis, que soy Catedrático)... además no hay ninguna diferencia: yo sigo escribiendo todos los papers... ¡Miserables! Eso sí, Ricardo los corrige bien, bien. En boli rojo, siempre.

En estos 16 años que mencionaba antes he llegado lejos. Nunca me imaginé que lo conseguiría tan pronto. Es imposible agradecer a todo el mundo que ha contribuido a esto, no acabaría nunca. Sí que me gustaría, sin embargo, mencionar a algunos a los que considero completamente imprescindibles. Mis Amigos: Antonio Portilla, Silvia Jiménez, Lucas Cuadra y Enrique Alexandre, de la UAH. Antonio Caamaño y Mihaela Chidean, de la URJC (Antonio, es la segunda vez que apareces en los agradecimientos de una Tesis Doctoral mía. Esto se está convirtiendo en una mala costumbre. Vas a ir al mar con un peso de 500 kilos.). César Hervás y Pedro Gutiérrez de la UCO. Carlos Casanova de AEMET. Luis Prieto, de Iberdrola. David Camacho de la UAM. Javier Del Ser y Sergio Gil, de Tecnalia. Gracias a todos por enseñarme,

y trabajar tanto y tan bien, a pesar de las dificultades.

Gracias a mis padres, Gregorio y Ana, a mi hermana, Susana, a mi cuñado y a mis sobrinos, Bruno y Olivia. Gracias a mi tía Mari y a mi tío Vitín. Gracias a mis suegros, cuñados, mi sobrino Rubén y a la tía Polo.

Y finalmente Gracias a lo mejor de mi vida: María Jesús y Darío.

Contents

List of Figures

List of Tables

Acronyms

2Px	2-points crossover
AFWA	Air Force Weather Agency (USA)
AI	Artificial Intelligence
AM	Analogue Method
ANFIS	Adaptive Neuro-fuzzy Inference System
ANN	Artificial Neural Network
AO	Antarctic Oscilation
ARMA	Auto-Regressive Moving Average model
ARIMA	Auto-Regressive Integrated Moving Average model
BSA	Backtracking Search Algorithm
CNN	Convolutional Neural Networks
CRO	Coral Reefs Optimization algorithm
CRO-SL	Coral Reefs Optimization with Substrate Layer algorithm
CRO-SP	Coral Reefs Optimization with Species algorithm
DECRO	Differential Evolution Coral Reefs Optimization algorithm
DNN	Deep Neural Network
DE	Differential Evolution
EA	Evolutionary Algorithm
ECMWF	European Centre for Medium-Range Weather Forecasts
ELM	Extreme Learning Machine
EMD	Empirical Mode Decomposition
ENSO	El Niño/Southern Oscilation
FAA	Federal Aviation Administration (USA)
FSL	Forecast Systems Laboratory (USA)
FSP	Feature Selection Problem
GA	Genetic Algorithm
GFS	Global Forecasting System
GGA	Grouping Genetic Algorithm
GP	Gaussian Process
GS	Gravitational Search algorithm
GSR	Global Solar Radiation
GHI	Global Horizontal Irradiance
HMCR	Harmony Memory Considering rate

HS	Harmony Search
k-NN	k-Nearest Neighbors
LA-CRO	Learning Automata Coral Reefs Optimization algorithm
LR	Linear Regression
MASL	Metres Above Sea Level
MCP	Measure Correlate Predict approaches
MERRA	Modern-Era Retrospective analysis for Research and Applications
ML	Machine Learning
MLP	Multi-Layer Perceptrons
MPx	Multi-points crossover
MLR	Multi-Linear Regression
MSE	Mean Squared Error
MSVR	Multi-output Support Vector Regression
NARX	Nonlinear Auto-Regressive eXogenous model
NCAR	National Center for Atmospheric Research (USA)
NCEP	National Center for Environmental Prediction (USA)
NDBC	National Data Buoy Center of the USA
NN	Neural Network
NOAA	National Oceanic and Atmospheric Administration (USA)
NWP	Numerical Weather Prediction models
NC	Neural Computation
PAR	Pitch Adjusting Rate
PCA	Principal Component Analysis
PSO	Particle Swarm Optimization
PV	Photo-Voltaic sytems
RBF	Radial Basis Function networks
RE	Renewable Energy
RF	Random Forest
RMSE	Root Mean Square Error
RS	Representative Selection
RSVR	Reduced Support Vector Regression
SA	Simulated Annealing
SM	Standard integer Mutation
SURFRAD	NOAA Surface Radiation Network (USA)
SVM	Support Vector Machine
SVR	Support Vector Regression
TOE	Tons Oil Equivalent
WEKA	Waikato Environment for Knowledge Analysis
WRF	Weather Research and Forecasting system

Chapter 1

Introduction

1.1 Motivation and objective

In the last decade, global energy demand has increased to non-previously seen levels, pushed by the increase in global population, fierce urbanization in developed countries and aggressive industrial development all around the world [Suganthi2012]. Conventional fossil-based energy sources have limited reservoirs and a deep environmental impact (contributing to global warming) [Bauer2015], and therefore they cannot satisfy this global demand for energy in a sustainable way [Rowland2016]. These issues related to fossil-based sources have led to a very important development of RE sources in the last years, mainly in renewable technologies such as wind or solar-based. The main problem with RE resources is their dependency on meteorological conditions, sometimes extreme, such as in the case of wind and solar energy [Kumar2016]. The fact that individual renewable sources cannot provide continuous power supply because of their uncertainty and intermittent nature is also an important issue which must be solved to improve the penetration of these resources in the electric system. Note that nowadays renewable energy resources cover approximately 19% of the total World energy demand [REN21-2017] (see Figure 1.1). This figure is still far from the non-renewable sources, which are estimated to cover almost 78% of the total demand, specially in developing countries. A huge amount of research is being conducted to obtain a higher penetration of renewable resources into the electric system.

Currently, the development of Information Technologies has led to a new era in computation, mainly ruled by *Data Science*. Data science and related technologies are nowadays of major importance in our society, since they play the highest roles in the development of a global, fully-connected, economy. In this regard, Soft-Computing and its related paradigms (Machine Learning, Computational Intelligence, etc.) have proven to be excellent tools to cope with difficult problems arisen in Data Science, in a huge variety of applications such as Medicine [Bisaso2017], Manufacturing [Sharp2018] or social networks [Bello2016], among many others. In this sense, Soft-Computing techniques have also been successfully applied to RE and atmospheric-related problems [Gardner1998, Renani2016, Kalogirou2006, Heinermann2016, Jha2017, Voyant2017,

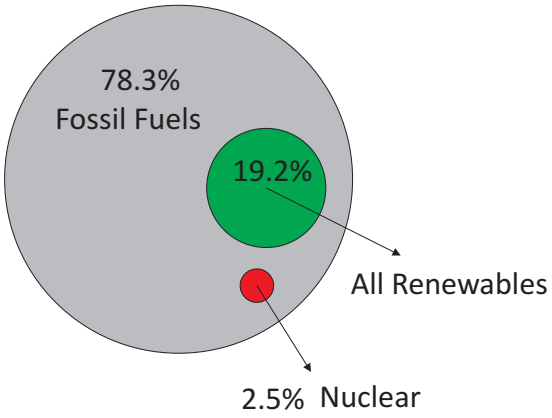


Figure 1.1: Implantation percentage of Renewables, Fossil Fuels and Nuclear energy at global level by 2016 [REN21-2017].

Sharma2018].

Soft-Computing is a branch of AI whose objective is to obtain robust problem solving techniques which emulate the human thinking. It is a huge research area, which merges different paradigms and lines of research including statistical learning, neural and cognitive signal processing, natural computing or fuzzy logic, among others. Moreover, Soft-Computing is a live research field, where new approaches are proposed for dealing with very hard problems which arise every day in Engineering and Science. Figure 1.2 shows a scheme of Soft-Computing which includes a first classification of techniques, such as Neural Computation and Learning Systems, dealing with regression and classification problems, or any type of supervised prediction problem, Evolutionary computation, a wide area which includes natural computing approaches, mainly devoted to optimization problems and systems, and finally Fuzzy systems, a line more linked with robotics and control systems.

This Ph.D. Thesis discusses new hybrid Soft-Computing approaches to solve problems related to RE and atmospheric science. We explore new approaches that merge a new type of Evolutionary Algorithm (the CRO algorithm), in different versions, with some kind of prediction or reconstruction approaches, such as NNs or the AM. The motivation of this work is the necessity for obtaining new powerful approaches which lead to better solutions in difficult problems (those that cannot be properly solved with traditional methods). The application of Soft-Computing techniques is fully justified in these cases, offering an alternative which is currently being exploited by many researchers and practitioners, with paradigms such as Big Data or Deep Learning [LeCun2015]. In this sense, hybrid approaches mixing Soft-Computing with alternative ML algorithms or methods is a line followed in the last years by many researchers [Tasci2014].

Next sections of this chapter will give a detailed description of the Soft-Computing algorithms

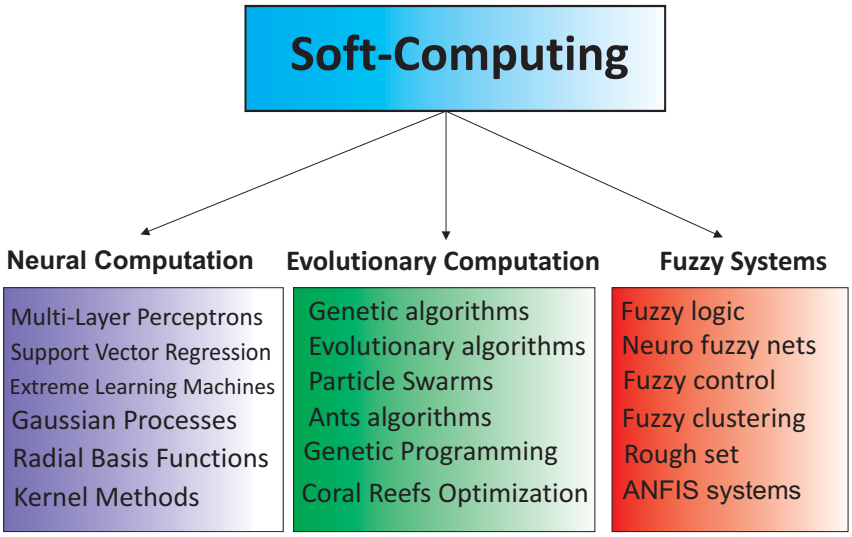


Figure 1.2: Soft-Computing sub-branches, including Neural Computation and Learning Systems, Evolutionary Computation and Fuzzy Systems.

proposed and alternative methods applied. At the end of the chapter, the structure of the rest of this Thesis will be discussed, where the main algorithms proposed and problems tackled are outlined.

1.2 Soft-Computing algorithms and methods used

1.2.1 Neural Computation-based approaches

Neural Computation is a part of Soft-Computing that includes algorithms inspired on how the human brain learns. It is based on algorithms usually known as ANNs. Neural Computation includes a large amount of different NNs, which have been mainly used in classification and regression problems. In this work, we consider *feed forward* neural networks, a type of neural processing approach which processes the information in different layers, with a privileged direction. MLPs are the most used feed forward neural approaches, and will be described next, together with ELMs, a kind of feed forward network with a very fast training scheme, perfect for constructing hybrid algorithms.

Multi-layer perceptrons

A MLP is a particular kind of ANN which is massively parallel. It is considered a distributed information-processing system, which has been successfully applied in modelling a large variety of nonlinear problems [Haykin1998, Bishop1995]. The MLP consists of an input layer, a number of hidden layers, and an output layer, all of which are basically composed by a number of special

processing units called *neurons*, as Figure 1.3 shows. As important as the processing units themselves is their connectivity, i.e. how the neurons within a given layer are connected to those of other layers by means of weighted links. These weight values are closely related to the learning ability of the MLP, and also with its ability to generalize the learning from enough number of examples. Thus, note that such a learning process demands a proper database containing a variety of input examples or patterns with the corresponding known outputs (tags). The adequate values of the neuron weights minimize the error between the output generated by the MLP (when fed with input patterns in the database), and the corresponding expected output in the database. The number of neurons in the hidden layer is a parameter to be optimized when using this type of neural network [Haykin1998, Bishop1995].

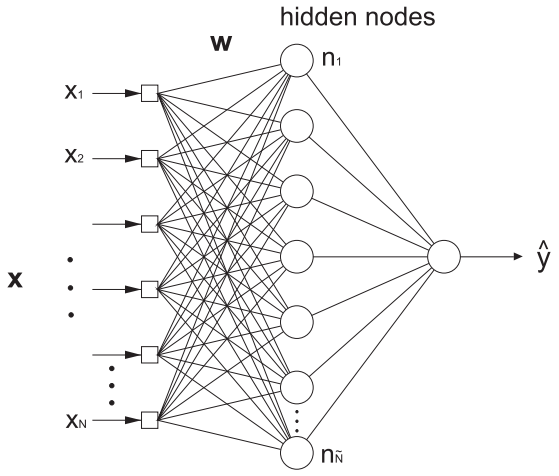


Figure 1.3: An scheme of Artificial Neural Network model.

The input data for the MLP consists of a number of samples arranged as input vectors, $\mathbf{x}=\{x_1,\ldots,x_N\}$. Once a MLP has been properly trained, validated and tested using an input vector different from those contained in the database, it is able to generate a proper output y. The relationship between the output and the input signals of a neuron is the following:

$$y = \varphi\left(\sum_{j=1}^{n} w_j x_j - \theta\right), \tag{1.1}$$

where y is the output signal, x_j for $j = 1,\ldots,n$ are the input signals, w_j is the weight associated with the j-th input, and θ is a threshold [Haykin1998, Bishop1995]. The transfer function φ is usually considered as the logistic function,

$$\varphi(x) = \frac{1}{1+e^{-x}}. \tag{1.2}$$

The process to obtain an accurate output is related to the training procedure as it was mentioned before. During the training process, the error between the estimated output and its

corresponding real value in the database will determine to what degree the weights in the network should be adjusted. Hence, the objective of the network training is to find the combination of weights which results in the smallest training error with the best possible generalization of the result. There are different algorithms that can be used to train a MLP. One possible technique is the back-propagation training algorithm [Bishop1995] which uses the procedure known as *gradient descent* to try to locate the global minimum of the error [Gardner1998]. Another approach is the well-known Levenberg-Marquardt [Hagan1994].

Extreme Learning Machine

An ELM [Huang2015, Huang2006] is a novel and fast training method based on the structure of MLPs, shown in Figure 1.3. The most significant characteristic of the ELM training is that it is carried out just by randomly setting the network weights, and then obtaining a pseudo-inverse of the hidden-layer output matrix. The advantages of this technique are its simplicity, which makes the training algorithm extremely fast, and also its outstanding performance when compared to alternative sequential learning methods, usually better than other established approaches such as classical MLPs. Moreover, the universal approximation capability of the ELM network, as well as its classification capability, have been already proven [Huang2012].

The ELM algorithm can be summarized as follows: given a training set $\mathbb{T} = (\mathbf{x}_i, \boldsymbol{y}_i)|\mathbf{x}_i \in \mathbb{R}^n, \boldsymbol{y}_i \in \mathbb{R}, i = 1, \cdots, l$, an activation function $g(x)$, which a sigmoidal function is usually used, and number of hidden nodes ($\tilde{N}$),

1. Randomly assign inputs weights $\mathbf{w}_i$ and bias b_i, $i = 1, \cdots, \tilde{N}$.

2. Calculate the hidden layer output matrix $\mathbf{H}$, defined as

$$\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1\mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}}\mathbf{x}_1 + b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(\mathbf{w}_1\mathbf{x}_l + b_1) & \cdots & g(\mathbf{w}_{\tilde{N}}\mathbf{x}_N + b_{\tilde{N}}) \end{bmatrix}_{l \times \tilde{N}} \tag{1.3}$$

3. Calculate the output weight vector β as

$$\beta = \mathbf{H}^{\dagger}\mathbf{y_t}, \tag{1.4}$$

   where $\mathbf{H}^{\dagger}$ stands for the Moore-Penrose inverse of matrix $\mathbf{H}$ [Huang2006], and $\mathbf{y_t}$ is the training output vector, $\mathbf{y_t} = [\mathbf{y_t}_1, \cdots, \mathbf{y_t}_l]^{\mathbb{T}}$.

Note that the number of hidden nodes ($\tilde{N}$) is a free parameter of the ELM training, and must be estimated for obtaining good results. Usually, scanning a range of $\tilde{N}$ values is the best solution. The Matlab extreme learning machine implementation by G. B. Huang, freely available at [ELM2018], is often considered for ELM implementation.

1.2.2 The Coral Reefs Optimization Algorithm

The CRO is a novel evolutionary-type meta-heuristic approach for optimization, recently developed in [Salcedo2014d], which is based on simulating the corals' reproduction and coral reefs' formation processes. The CRO algorithm tackles optimization problems by modeling and simulating all the distinct processes in a real coral reef. Let $\mathcal{R}$ be a model of reef, consisting of a $N \times M$ square grid. We assume that each square (i, j) of $\mathcal{R}$ is able to allocate a coral (or colony of corals) $\Xi_{i,j}$, representing different solutions to our problem, using a given encoding for the problem at hand (an encoding is a representation of the optimization problem's variables, usually in the form of a chain of numbers, leading to a given *search space*). The CRO algorithm is first initialized at random by assigning some squares in $\mathcal{R}$ to be occupied by corals (i.e. solutions to the problem) and some other squares in the grid to be empty, i.e. holes in the reef where new corals can freely settle and grow. The rate between free/occupied squares in $\mathcal{R}$ at the beginning of the algorithm is a parameter of the CRO algorithm denoted as ρ, and note that $0 < \rho_0 < 1$.

After the reef initialization, a second phase of reproduction and reef formation is carried out. First, a simulation of the corals' reproduction in the reef is done by sequentially applying different operators for modeling sexual reproduction (broadcast spawning and brooding), asexual reproduction (budding), and polyps depredation:

1. **Broadcast Spawning (external sexual reproduction):** the modeling of coral reproduction by *broadcast spawning* consists of the following steps:

   1.a. In a given step k of the CRO algorithm, select uniformly at random a fraction of the existing corals ρ_k in the reef to be broadcast spawners. The fraction of broadcast spawners with respect to the overall amount of existing corals in the reef will be denoted as F_b. Corals that are not selected to be broadcast spawners (i.e. $1 - F_b$) will reproduce by brooding later on in the algorithm.

   1.b. Select couples out of the pool of broadcast spawner corals in step k. Each of such couples will form a coral larva by means of a given crossover mechanism or any other exploration strategy. Note that once two corals have been selected to be the parents of a larva, they are not chosen anymore in step k (i.e. two corals are parents only once in a given step). This couple selection can be done uniformly at random or by resorting to any fitness proportionate selection approach (e.g. roulette wheel).

2. **Brooding (internal sexual reproduction):** at each step k of the reef formation phase in the CRO algorithm, the fraction of corals that will reproduce by brooding is $1 - F_b$. The brooding modeling consists of the formation of a coral larva by means of any kind of mutation mechanism, in order to simulate of the brooding-reproductive coral (self-fertilization considering hermaphrodite corals).

3. **Larvae setting:** once all the larvae are formed at step k either through broadcast spawning (1.) or by brooding (2.), they will try to set and grow in the reef. First, the health function (fitness) of each coral larva is computed. Second, each larva will randomly try to set in a square (i, j) of the reef. If the square is empty (free space in the reef), the coral grows therein no matter the value of its health function. By contrast, if a coral is already occupying the square at hand, the new larva will set only if its health function is better than that of the existing coral. We define a number κ of attempts for a larva to set in the reef: after κ unsuccessful tries, it is considered as depredated by the animals in the reef.

4. **Asexual reproduction:** the modeling of asexual reproduction (budding or fragmentation) is the CRO is carried out in the following way: the overall set of existing corals in the reef are sorted as a function of their level of health (given by $f(\Xi_{ij})$). Then a small fraction F_a duplicate themselves and are mutated in order to obtain variability. These new corals try to settle in a different part of the reef by following the setting process described in Step 3.

5. **Depredation in polyp phase:** corals may die during the reef formation phase of the CRO algorithm. At the end of each reproduction step k, a small number of corals in the reef can be depredated, thus liberating space in the reef for next coral generation. The depredation operator is applied with a very small probability P_d at each step k, and exclusively to a fraction F_d of the worse health corals in $\mathcal{R}$.

Figure 1.4 illustrates the flowchart diagram of the CRO algorithm, with the different CRO phases (reef initialization and reef formation), along with all the operators described above.

1.2.3 Advanced CRO models

The basic CRO can be improved to obtain stronger versions of the meta-heuristic, based on alternative processes that occur in coral reefs. We describe here two different modifications of the CRO algorithm which improve its performance in specific applications. First, we describe the CRO with species, which helps tackle optimization problems with variable length encodings. It is also useful for managing different encodings of problems within the same population, obtaining a competitive co-evolution algorithm. The second version is the CRO with substrates layer. It has been useful to obtain a competitive co-evolution algorithm in which different models are applied to the same problem. These models can be either exploration models, repairing mechanisms, etc., and the only pre-requisite is that the objective function to evaluate corals in the reef must be the same for the different models considered.

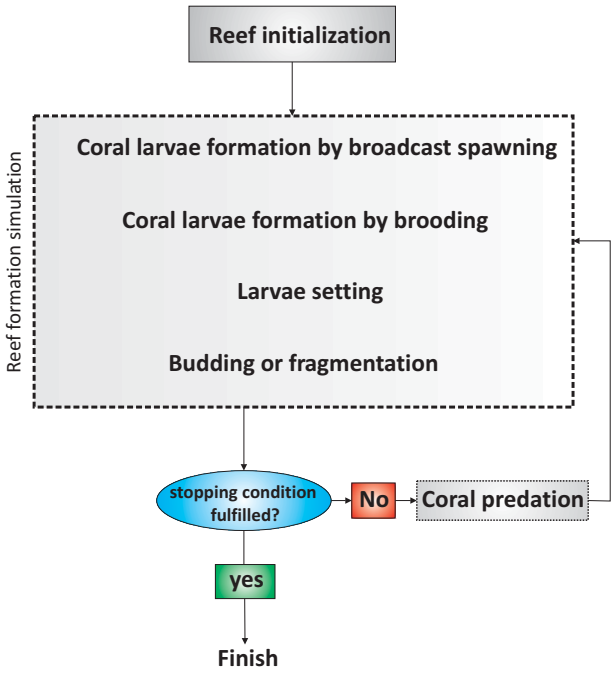


Figure 1.4: Flowchart diagram of the original CRO algorithm.

CRO with species

The first modification of the CRO consists in considering different coral species within a single coral community (CRO-SP). The objective of this modification is that each coral species represents a different model (or its hyper-parameters) out of T possible models. In this context, *model* is generic, so it can represent either a different encoding for the problem, a different way of calculating the objective function, etc. The CRO-SP is a new powerful way of managing optimization problems with variable or different encodings. In this case, each species will represent a different encoding, and the idea is that only corals of the same species can reproduce in the broadcast spawning operator. Note however, that all the models compete together in the larvae setting, since the objective function in all cases should be the same for all the species.

The CRO-SP was first introduced in [Salcedo2017b] as a methodology to deal with a Model Selection Problem, in an application of total energy consumption prediction in Spain. In [Salcedo2017b] each species represents a different way of calculating the total energy consumption (a different model), and the idea was to build a competitive co-evolution approach that obtained the best possible model in addition to alternative parameters such as the best prediction variables to feed the prediction model. Note that the CRO-SP could be also used to evolve a competition of different regressions for a given problem, for example neural networks, support vector machines, etc., in which the CRO encodes the parameters of each regressor.

The following algorithm shows an outline of the CRO containing multiple species. Note

that the competition among species will produce emerging behavior, so the best model (species) eventually will dominate, and will occupy the majority of spaces in the reef.

Require: Valid values for CRO parameters.
Ensure: The best model out of T possible.
1: Algorithm initialization (T different species)
2: **for** each iteration of the CRO **do**
3: Update values of influential variables: mortality probability and the probability of asexual reproduction
4: Asexual reproduction (budding or fragmentation)
5: Sexual reproduction 1 (broadcast spawning, only same species can reproduce)
6: Sexual reproduction 2 (brooding, only same species can reproduce)
7: Settlement of new larvae (competition among species)
8: Mortality process
9: Evaluate the new population in the reef (with the specific model given for each species)
10: **end for**

CRO with Substrate Layers

The second important modification of the CRO is the incorporation of substrate layers (CRO-SL). It is based on the fact that there are many more interactions in real reef ecosystems which can be also modelled and incorporated to the CRO approach to improve it. For example, different studies have shown that successful recruitment in coral reefs (i.e., successful settlement and subsequent survival of larvae) depends on the type of substrate on which they fall after the reproduction process [Vermeij2005]. This specific characteristic of the coral reefs was first included in the CRO in [Salcedo2017b], in order to solve different instances of the Model Type Selection Problem for energy applications. In [Salcedo2016c, Salcedo2017b], different substrate layers were defined in the CRO, in such a way that each layer represents a different model to evaluate the energy demand estimation in Spain, from macro-economic variables.

As in the case of the CRO-SP, the CRO-SL is a very general approach: it can be defined as an algorithm for competitive co-evolution, where each substrate layer represents different processes (different models, operators, parameters, constraints, repairing functions, etc.). The inclusion of substrate layers in the CRO can be done in a straightforward manner: we redefine the artificial reef considered in the CRO in such a way that each cell of the reef $\mathcal{R}$ is now defined by 3 indices (i, j, t), where i and j stand for the cell location in the grid, and index $t \in T$ defines the substrate layer, by indicating which structure (model, operator, parameter, etc.) is associated with the cell (i, j). Each coral in the reef is then processed in a different way depending on the specific substrate layer in which it falls after the reproduction process. Note that this modification of the basic algorithm does not imply any change in the corals' encoding (all the corals in the algorithm are encoded in the same way).

The CRO-SL has been applied in [Salcedo2016c] to obtain a competitive co-evolution algorithm in which each substrate is assigned to a different implementation of an exploration procedure. Thus, each coral will be processed in a different way in the reproduction step of the algorithm depending on the substrate it occupies. Figure 1.5 shows an example of the CRO-SL, with different substrate layers. Each one is assigned to a different exploration process, Harmony Search based, Differential Evolution, 1-point crossover or Gaussian mutation (alternative assignment and different exploration processes can be used in the substrate layer of the CRO-SL approach).

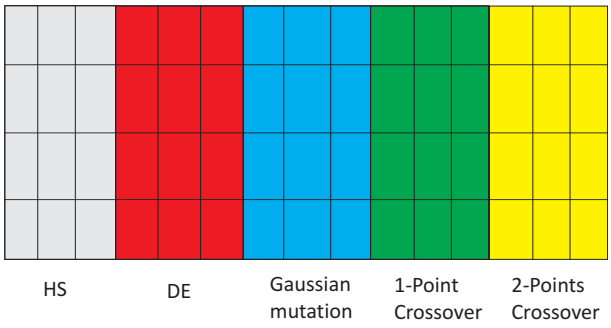


Figure 1.5: Example of reef in the CRO-SL, where there are five substrate layers associated with the broadcast spawning process. Each substrate layer represents now a different exploration process to carry out in that substrate.

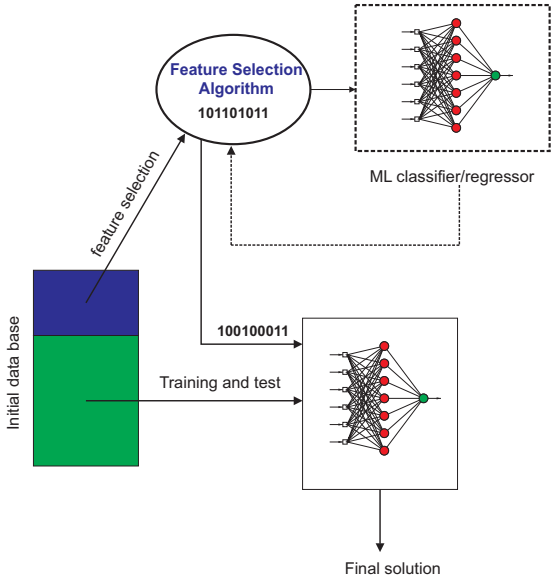
The CRO-SL is a general procedure to co-evolve different models, operators, parameter values, etc., with the only requisite that there is only one health function defined in the algorithm. In other words, each substrate can include a different processing of problem's constraints, exploration or exploitation procedures etc.

1.3 Feature Selection in Machine Learning

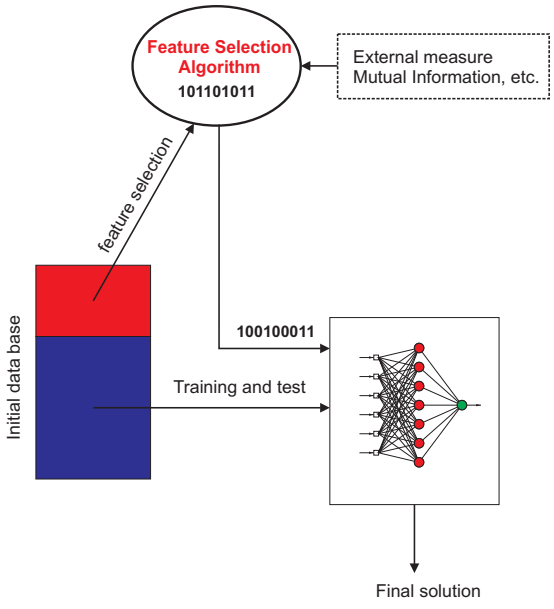
Feature selection is an important task in ML-related problems because irrelevant features, used as part of the training procedure of different prediction systems, can increase the cost and running time of the system, and make its generalization performance much poorer [Blum1997, Weston2000]. In its more general form, the FSP for a learning problem from data can be defined as follows: given a set of labeled data samples $(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_l, y_l)$, where $\mathbf{x}_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$ (or $y_i \in \{\pm 1\}$ in the case of classification problems), choose a subset of m features ($m < n$), that achieves the lowest error in the prediction of the variable y_i.

Thus, note that there are many different algorithms which can be used to solve a FSP. In general, they can be structured in two different families or paradigms:

- The *wrapper approach* to the FSP was introduced in [John1994]. The feature selection algo-



(a)



(b)

Figure 1.6: (a) Outline of a Wrapper method; (b) Outline of a Filter method.

rithm conducts a search for a good subset of features using the classifier/regressor itself as

part of the evaluating function. Figure 1.6 (a) shows the idea behind the wrapper approach: the classifier/regression technique is run on the training dataset with different subsets of features. The one which produces the lowest estimated error in an independent but representative test set is chosen as the final feature set. For further reading on wrappers methods, the following classical works can be consulted [Kohavi1997, Yang1998, Salcedo2002]. In the case of the wrapper method, the FSP admits a mathematical definition as follows: The FSP consists of finding the optimum n-column vector $\boldsymbol{\sigma}$, where $\sigma_i \in \{1, 0\}$, that defines the subset of selected features, which is found as

$$\boldsymbol{\sigma}^o = arg \min_{\boldsymbol{\sigma},\boldsymbol{\alpha}} \left(\int V(y, f(\mathbf{x} * \boldsymbol{\sigma}, \boldsymbol{\alpha})) dP(\mathbf{x}, y) \right), \tag{1.5}$$

where $V(\cdot, \cdot)$ is a loss functional, $P(\mathbf{x}, y)$ is the unknown probability function the data was sampled from and we have defined $\mathbf{x} * \boldsymbol{\sigma} = (x_1\sigma_1, \ldots, x_n\sigma_n)$. The function $y = f(\mathbf{x}, \boldsymbol{\alpha})$ is the classification/regression engine that is evaluated for each subset selection, $\boldsymbol{\sigma}$, and for each set of its hyper-parameters, $\boldsymbol{\alpha}$.

- In the *filter approach* to the FSP, the feature selection is performed based on the data, ignoring the classifier algorithm. An external measure calculated from the data must be defined to select a subset of features. After the search, the best feature subset is evaluated on the data by means of the classifier/regression algorithm. Note that filter algorithms performance completely depends on the measure selected for comparing subsets. Figure 1.6 (b) shows an example of how a filter algorithm works. Filter methods are usually faster than wrapper methods. However, their main drawback is that they totally ignore the effect of the selected feature subset on the performance of the classification/regression algorithm during the search. So, usually their performance is poorer than wrapper approaches. Further analysis and application of filter methods can be found in [Torkkola2000, Torkkola2002].

- There are different works which have combined both wrapper and filter methodologies to build hybrid approaches. They have shown good performance in specific applications [Ferreira2014, Huda2014, Huda2016, Solorio2016].

For both wrapper and filter methods, a binary representation can be used for the FSP, where a 1 in the i_{th} position of the binary vector means that the feature i is considered within the subset of features, and a 0 in the j_{th} position means that feature j is not considered within the subset. Note that using this notation is equivalent to encode the problem as the vector $\boldsymbol{\sigma}$ included in expression (1.5). Note also that there are 2^n different subsets (where n is the total number of features), and the problem is to select the best one in terms of a certain measure, which can be either internal (wrapper methods) or external (filter methods) to the classifier/regressor. Alternative encodings, such as integer vectors, are however also possible, and even more adequate in some specific applications.

1.4 Representative measuring points selection

In ML, a *Representative Selection* (RS) problem consists of finding exemplar samples from a given data, points or items collection, in such a way that the selected exemplars accurately summarize the complete set of starting data [Wang2017]. RS problems appear in many different Science and Engineering problems, such as Computer Science [Wang2017], Bio-mechanics [Katrin2012] or Networking [Poorter2017]. In Climatology, RS problems have been faced in different applications, such as dynamical climate downscaling [Rife2013], selecting regional climate scenarios [Wilcke2016], selecting the most representative subset of global climate models in terms of a given error measure [Ruane2017], selecting the most representative models for climate change studies [Lutz2016], or optimizing the position of weather monitoring stations [Amorim2012].

Let $F(\mathbf{s}, t)$ be a field of a climatological variable, defined in a set of $|S|$ measuring points or stations $\mathbf{s}$, during an observation period $\hat{t}$. Such period can be split into a training period, t^T, and a validation or test period of duration t^V, in such a way that $\hat{t} = t^T + t^V$. Let χ be a given reconstruction algorithm for the field $F(\mathbf{s}, t)$. χ operates in a subset of measuring stations $\mathbf{s}_N$, where N ($N < |S|$) represents the number of measuring points or stations selected. Moreover, note that χ only uses the data in the training period to obtain the best reconstruction of the initial field $F(\mathbf{s}, t)$, in terms of an error measure $e(\mathbf{s}_N)$ evaluated in the test period. The RS problem we face in this Thesis consists of obtaining the best possible subset of N measuring points $\mathbf{s}_N^*$, which minimizes $e(\mathbf{s}_N^*)$ (usually e stands for a mean square or absolute error function). Note that the subset $\mathbf{s}_N^*$ stands for the N most representative measuring stations for the field $F(\mathbf{s}, t)$, in terms of the reconstruction algorithm χ considered. We have chosen the well-known Analogue Method as the reconstruction algorithm χ.

The Analogue Method in RS

The AM is a prediction/reconstruction algorithm very used in atmospheric sciences. It is based on the principle that two similar states of the atmosphere lead to similar local effects [Lorentz1969]. More specifically, two states of the atmosphere are considered as "analogues", when there is a resemblance between them, in terms of an analogy criterion and objective variables. Thus, the AM consists of searching for a certain number of past situations in a meteorological archive, in such a way that they present similar properties to that of a target situation for any chosen predictors or variables.

Different versions of the AM can be found in the literature, with a wide range of meteorological and climatological applications. In [Gibergans2007] the AM method is improved by using local thermodynamic data to predict autumn precipitation in Catalonia, Spain. In [Chardon2014], the AM was applied to downscale precipitation in France, including the novelty of spatial similarity in the algorithm. In [DMonache2011] the AM was improved with a Kalman filter in order to improve numerical weather predictions. In [Yiou2014] the AM was applied to a palaeo-climatic problem, specifically, the meteorological reconstruction using circulation ana-

logues in the late Eighteen Century. In [Lguensat2017] the analog data assimilation technique was presented. It combines the AM with ML techniques based on the k-NN to obtain the most important data to be assimilated for numerical prediction systems. AM ensembles have also been applied to different forecasting problems. One of the original works proposing AM ensembles is [DMonache2013], where the method and its main characteristics have been presented and compared to alternative state-of-the-art numerical weather prediction ensemble systems. In [Junk2015] the statistics of the AM ensemble model are fully described. The AM ensemble has also been applied to specific prediction problems, such as in [Alessandrini2015], where an analog ensemble has been applied to probabilistic solar power forecast in three solar farms in Italy. A comparison with a quantile regression algorithm and persistence ensembles has proven the goodness of the AM approach in this problem. Another work dealing with AM ensembles is [Alessandrini2015b], with application to short-term wind power prediction. In this case the AM ensemble has been applied to the wind power production prediction of a wind farm in northern Sicily (Italy), comparing the performance with alternative algorithms such as quantile regression and numerical weather models. In [Vanvyve2015] another AM ensemble approach is presented, with application in wind resource estimation. The paper analyzes the wind resource of different locations in Europe and USA by applying an AM ensemble.

Finally, note that very recently the AM has been mixed with meta-heuristics algorithms. In [Horton2017] a genetic algorithm has been used to tune the parameters of an AM. The accuracy of this evolutionary-AM approach has been shown in a case study of probabilistic precipitation forecast in Switzerland.

In this Thesis we apply the AM to the RS problem, as follows: Given a subset of measuring points of stations $\mathbf{s}_N$, the AM process starts by obtaining the most similar situations in the past for the field $F(\mathbf{s}_N, t^V)$ (considering all the evaluation period). In other words, this is equivalent to, for each time $T \in t^V$, obtaining the most similar situation (or average of k most similar situations) in the past (training period), located in time $T^* \in t^T$, considering only the selected measuring points $\mathbf{s}_N$ (note that T^* depends on the T considered, i.e. $T^* = T^*(T)$). Then, the complete reconstruction of the field F is calculated by using the past situation $F(\mathbf{s}, T^*)$ and the objective situation $F(\mathbf{s}, T)$, and a reconstruction error is obtained. The final function $e(\mathbf{s}_N)$ is calculated as the root mean square error (RMSE) in the field reconstruction:

$$e(\mathbf{s}_N) = \sqrt{\frac{1}{|S| \cdot t^V} \sum_{T=1}^{t^V} \sum_{k=1}^{|S|} \left(F(s_k, T^*) - F(s_k, T)\right)^2} \tag{1.6}$$

Figure 1.7 graphically shows the process for the AM application (χ field reconstruction algorithm), and the calculation of the error function (RMSE) associated with $\mathbf{s}_N$.

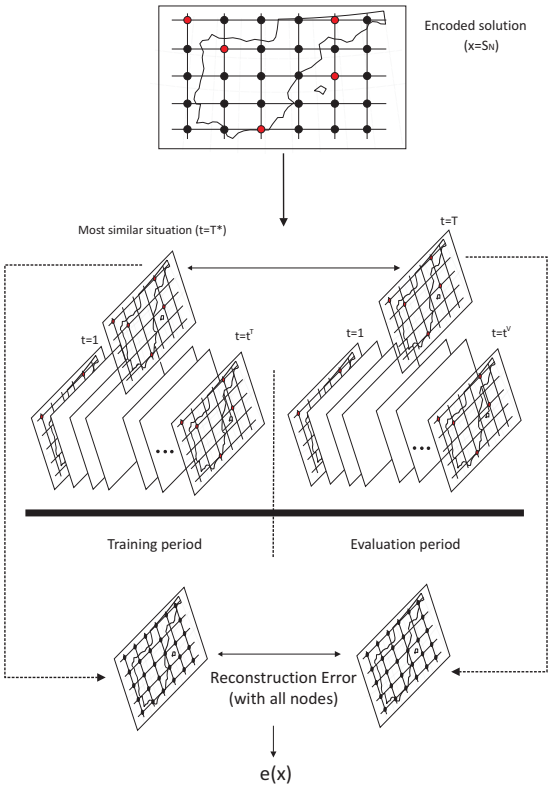


Figure 1.7: AM calculation process. From a selection of measuring points to be evaluated (marked in red), the AM approach works by calculating the field reconstruction error for all the evaluation period, by looking for the most similar situation in the training period, considering only the information provided by the selected measuring points.

1.5 Structure of the Thesis

The rest of this Thesis is organized in a state-of-the-art chapter, where previous works related to the Thesis content are reviewed and analyzed, and two different technical parts:

1. First, contributions with results in RE-related problems are described, which include new hybrid algorithms involving different CRO versions and ELM for problems of feature selection in wind and solar resource prediction problems.

2. Second, results in two different RS problems are discussed: optimal measuring points selection for Temperature and Wind fields reconstruction with a hybrid CRO-SL algorithm and the AM reconstruction approach.

To conclude, some final remarks and future research lines are summarized in the last part of the document, with a list of publications produced by the research carried out in this Thesis

shown in a final Appendix section.

Chapter 2

Methods and algorithms: Previous work

This section presents a description of the state-of-the-art in the technological fields addressed in this Thesis. The idea of this chapter is to focus on the discussion of previous works dealing with the techniques applied in the Thesis.. First, a review of Feature Selection in renewable energy systems is carried out. Wind and solar resources short-term forecasting are first considered, though references in other energy-related problems have also been reviewed. A review of literature devoted to long-term wind speed variability is then discussed. The section is closed with a review of previous work dealing with the CRO algorithm and its variants, with special focus on the results of the algorithm in renewable energy problems.

2.1 A review of FSP methods in renewable energy prediction problems

This section presents a review of the main works dealing with FSP in renewable energy prediction problems. The section is structured in FSP in wind energy, FSP in solar energy and FSP in energy-related problems.

2.1.1 Feature selection in wind energy prediction systems

Wind energy is the most developed renewable energy technology, and, thus, research on its prediction has been carried out during more than 20 years. FSPs have also been tackled in wind energy during more than a decade. For example, Jursa in [Jursa2007], first proposed a feature selection in wind power prediction systems. Specifically, a PSO approach was introduced in order to obtain the optimal set of features which provided the best prediction. That work was further completed in another paper by Jursa and Rohrig, [Jursa2008], where NN and k-NN algorithms were used as prediction models, and two wrapper feature selection models were

considered. A large number of features were used in that work as inputs for the prediction methods, such as previous wind speed and power samples (historic data) and predicted weather variables. Two different global search mechanisms were used to form the wrapper approaches: a PSO algorithm and a DE approach. The encoding of the FSP into the PSO and DE algorithms consisted of different values of the predictive variables delayed in time, until a maximum delay d was reached. The evaluation of the systems was carried out in 10 wind farms in northwestern Germany, obtaining good performance in the prediction of the wind power in each wind farm. A comparison of the wrapper PSO and DE algorithms with a NN and k-NN algorithms was carried out against a manually constructed solution, which included values of wind speed, wind power and a cyclic time series to capture the intra-daily variability. Persistence was also included as a baseline (reference) algorithm. It was clearly shown that the results obtained by the wrapper approaches outperformed the manually constructed features, and also the persistence algorithm in the 10 wind farms considered. Improvements around 5% (depending on the wind farm considered) were obtained when using the PSO and DE algorithms for feature selection versus the neural network working without feature selection mechanism. The improvement of including feature selection against persistence is higher, reaching values near 15% depending on the wind farm studied, with a mean improvement over 10% with the hybrid PSO-NN algorithm.

Gupta et al. [Gupta2011], proposed a hybrid wrapper approach, formed by a GA and a NN as predictor. The database consisted of 12 different variables, which were measured in a wind farm in Jaipur, India. The input variables were a combination of atmospheric variables and previous wind speed measurements. The system for feature selection improved the NN with all the variables as inputs for the prediction, obtaining an improvement over 5% in prediction accuracy. The authors show that the best result was obtained with 8 features out of the total 12. Results showed an improvement up to 5% when the feature selection was included in the system. Note that, in this case, the number of predictive variables involved in the problem was small, so the use of a meta-heuristic to carry out the search was not necessary. In fact, given a partition of the search space and fixing all the NNs parameters, the best solution would be found by examining all the possible combinations of features ($2^{12}-1$, since the 0 was discarded). Note that the use of meta-heuristic approaches is recommended for FSP problems where the search space is extremely large.

More advanced computational methods have been recently applied to different feature selection problems in wind energy estimation. In [Kou2014], Kou et al. proposed an online feature selection method with a wrapped GP as predictor, for a problem of probabilistic wind power forecasting. Experiments in real data from several wind farms in the Baotou region, China, showed the performance of the proposed methodology. Short term wind speed prediction (15 minutes and 1 hour time-horizon predictions) was considered. In this case, a large number of predictive variables from direct measurements in the wind farm and also from numerical weather modes were taken into account. Regarding observational variables, this work considered 384 inputs, from 4 measurement points, including data until four days before the prediction. The input variables from the numerical weather model were obtained from the GFS system (NOAA), with

a resolution of 2.5 degrees. These variables were linearly projected to the four measurement points. The numerical variables considered to improve the performance of the prediction system were measurements of 10m wind speed. The feature selection was carried out by means of a greedy algorithm, the sequential forward greedy search, which starts from an empty set of variables, then incorporates selected features one by one, and finally evaluates the set. The algorithm keeps the best set of features found in the search. Note that this search is not as effective as applying a meta-heuristic approach, in which different sets of features are evaluated together. However, it is a faster procedure, and ensures a good enough solution with limited computational resources. The results obtained in the paper showed improvements of 6% with respect to persistence, and smaller improvements (less than 1%) versus non-linear systems such as NNs or SVRs.

In [Salcedo2015c], Salcedo et al. proposed a novel hybrid FSP wrapper approach formed by a CRO with HS operators for feature selection with an ELM as regressor, for wind speed prediction in wind farms. A large number of features were considered: temperature, wind speed and direction at different heights, etc. The predictive variables were obtained from numerical weather models in the surroundings of the wind farm, and also included synthetically constructed variables from the original ones. The final wind speed prediction process after the feature selection was also carried out by using an ELM. In this work, the evaluation of the CROHS-ELM approach was carried out in real data from a wind farm in Oregon, USA, and compared to a different wrapper approach guided by an evolutionary algorithm. The results showed the goodness of the proposed CRO-ELM approach in comparison to the EA-ELM, with an improvement around 4%.

In [Carta2015], Carta et al. analyzed different feature selection methods within NNs as MCP approaches. MCP tries to predict the wind speed at a given point taking into account measurements in alternative (usually closely situated) sites. Thus, the features (input data) involved in the prediction problem are usually wind speed and direction in neighbour locations. In this case, filter and wrapper methods were tested as feature selection mechanisms, in a problem of mean hourly wind speeds and directions recorded in 2003 and 2004 at five weather stations in the Canary Islands. As mentioned, the input data were the wind speed and direction at different previous times in 4 measurement stations, and the objective was to obtain the wind speed prediction in the objective station. This work discussed the performance of a filter method for feature selection, based on correlation between input variables. In this case, the features selection worked by ranking the variables depending on their linear correlation coefficient, and keeping the less correlated ones in order to serve as inputs for the NN. In a second round of experiments, a wrapper approach for feature selection was investigated by applying an exhaustive search, which used the NN to provide the accuracy of each features set. A comparison with the case of no feature selection was carried out. The results showed that the feature selection mechanisms led to an effective improvement. They were analyzed in terms of statistical tests, where it was shown that the feature selection methods are more effective when there is a low correlation between the objective and the input measurements sites. On the other hand, the statistical tests showed

no improvements after applying feature selection if the objective measurement tower is highly correlated with the predictive towers.

In [Kong2015], Kong et al. presented a wind speed prediction system based on a Support Vector Regression approach. A modified version of the algorithm which selects a subset of data as support vectors and solves a smaller optimization problem (RSVR) was used. A problem of very short-term wind speed prediction (few minutes time-horizon) was tackled in this case. Wind speed and direction, temperature and pressure were the input variables. A PCA of the data was used as a feature extraction method to determine the most important factors affecting the wind speed. Note that, in this case, the SVR works on a different space (given by the PCA), so the initial input space should be transformed previously to the prediction. In general, feature extraction is possible, but less popular than feature selection, since the importance of each real feature cannot be evaluated using, for example, the PCA transform. This is a general issue with feature extraction methods, which impedes the physical interpretation of the features. The evaluation of the proposed RSVR was carried out in data from a real wind power plant located in Neimenggu, China. The experiments consisted of comparing the performance of the RSVR with PCA against those by the SVR and a RBF network. It was shown that the performance of the RSVR with PCA is better than that of the comparing algorithms, improving over 8% the performance of the best counterpart.

Kumar and López, in [Kumar2015], proposed a NARX approach to tackle a problem of wind speed prediction from meteorological series as input variables. NARX is a recurrent dynamic network, with feedback connections enclosing several layers of the network [Lin1996]. It has performed better than alternative NNs in highly non-linear problems. A feature selection was carried out previously to the prediction step by applying the ReliefF filter method. This method is a filter approach for feature selection based on estimating the quality of attributes according to how well their values distinguish between instances close to each other. In this case, a very short-term wind speed prediction problem from previously measured values (inputs) was tackled. Results in data from a meteorological tower at the University of Waterloo, Canada, were shown, though there was not a comparison with the performance of an alternative similar algorithm, which difficults its assessment. The results obtained showed that the feature selection applied was able to significatively improve the performance of the NARX algorithm over 20%, compared to the case when no feature selection algorithm was considered.

In [Zhang2016], Zhang et al. proposed two hybrid models which combine EMD, a feature selection process and machine learning regressors (NNs and SVRs), for a problem of short-term wind speed prediction. The idea is decomposing the original wind speed time series into a set of sub-series by applying the EMD method. Then, a feature selection process is introduced to identify the relevant and informative features for all the sub-series. A linear regression method completes this process, in which different subsets of features (sub-series serving as inputs for the wind speed estimation) were evaluated. The system was finally completed by applying a predictive algorithm (NN or SVR) to the final set of features obtained from the previous step. The performance of this system has been evaluated using real data from three different wind

farms in China, located at Jiangsu, Ningxia and Yunnan. The complete system with EMD and feature selection was compared to the performance of the NN and SVR approaches, showing the advantages of the proposed approach, with improvements up to 34% when the EMD and feature selection was considered. Another work dealing with EMD and feature selection is due to Jiang and Huang [Jiang2017], who presented a hybrid ensemble EMD approach with feature selection and error correction for a problem of short-term wind speed prediction. This system starts by decomposing the wind speed time series into a number of subseries by applying the EMD algorithm. Two feature selection mechanisms based on filter methods are then applied, including kernel density estimation-based Kullback-Leibler divergence and energy measure. After this process, two different SVR algorithms and an auto-regressive model were used to obtain the wind speed prediction, and also to correct the resulting error whether possible. Two different experiments in wind speed data from Colorado and Minnesota (USA) showed the effectiveness of this approach when compared with its different components on their own.

In [Zheng2017], Zheng et al. merged PSO, GS and a ELM in a single hybrid algorithm for wind speed prediction. The meta-heuristics algorithms were applied in two different ways: a real version of both PSO and GS approaches were used to tune the parameters of the ELM, whereas a binary version of both algorithms tackled a FSP related to the wind prediction. The approach was tested in a now-casting (10 min. prediction time-horizon) wind speed prediction problem at two locations of the NREL. A similar approach is presented by Zhang et al. in [Zhang2017], where also a hybrid model ELM with different meta-heuristics is built. In this case, the authors use a BSA hybridized with an ELM. A real-valued version of the BSA is exploited to search for the optimal combination of weights and bias of the ELM, while a binary version of the algorithm is applied to obtain the best set of features for the prediction system. This hybrid model was tested in two different wind speed prediction problems: a half-hour wind speed observation data from two wind farms in Inner Mongolia (China), and now-casting (10-min. wind speed data) from the Sotavento Galicia wind farm.

Finally, Feng et al. in [Feng2017], developed a multi-model methodology based on the combination of machine learning techniques for a problem of short-term wind speed prediction. NN, SVR, RF and Boosting machine were applied as an ensemble with a previous step of feature selection. A PCA together with a Granger causality test, auto-correlation analysis and recursive feature elimination were sequentially applied to reduce the number of features involved in the wind speed prediction. The proposed ensemble was able to provide a good quality wind speed prediction, including confidence intervals for the prediction. Comparison with alternative machine learning approaches without feature selection showed the impact of the FSP in the wind speed prediction at seven locations of the SURFRAD network (USA).

2.1.2 Feature selection in solar energy prediction systems

In [Fu2013], Fu and Cheng proposed a system for solar irradiance very short-term prediction (minutes time-horizon). The work used a solar irradiance prediction scheme with features ex-

tracted from all-sky images. The idea was to obtain proper features from all-sky images derived from an all-sky camera located in Taiwan, apply a feature extraction algorithm to the images, and then use a regression technique to predict a clearness index from them. In a second step, the clearness index was used to calculate the desired solar irradiance together with the extraterrestrial solar irradiance value, which only depends on astronomical variables. The results in real data from all-sky images in Taiwan collected in August 2011 showed that the proposed system was an accurate tool to estimate short-term solar irradiation prediction locally. The experimental results showed an improvement in the short-term prediction of solar irradiance of about 4% in comparison to the estimation of the solar irradiance directly from weather variables.

In [Yadav2014b], Yadav et al. carried out a study of the main influencing input parameters for solar radiation prediction with NNs in different locations of India, by using WEKA software. Different variables such as daily average temperature, minimum temperature, maximum temperature, altitude, sunshine hours and site location were considered. The study used a previous wrapper method with a regression tree implemented in WEKA to select the best set of features. After this process, several NNs implementations from WEKA were evaluated on this best set of features. Improvements over 13% were obtained after the FSP process (comparing the NNs without feature selection pre-processing) in some of the locations considered. Note that also in this case, the small number of features involved in the problems made an exhaustive search algorithm possible, which was implemented in WEKA software.

Wang et al. [Wang2015] applied a feature extraction method to select the best set of input parameters for a SVM classifier, in order to reconstruct a database of weather types. These weather types are directly connected with the photovoltaic power generation accuracy, so the most useful features were extracted from photovoltaic data, and they served as inputs for a SVM to reconstruct the database of weather types. This is a classification problem used as a previous step in the estimation of photovoltaic power generation prediction for buildings. Alternative classification problems and algorithms in renewable energy are described in [PerezO2016].

In [Rana2016], Rana et al. forecasted the electricity power generation by a solar photovoltaic system. Short-term prediction (from 5 to 60 minutes ahead) was considered. Input variables at different previous times were considered: solar irradiance, temperature, humidity and wind speed. A total of 4200 variables were finally available as inputs to predict the photovoltaic generation in the next hour, in intervals of 5 minutes. In this case the correlation-based feature selection, which selects the best set of variables with higher correlation with the objective variable, was used as filter method. After this feature selection, two machine learning algorithms were applied to generate the system's prediction: an ensemble of NNs and a SVR approach. Experiments in power data collected from the St. Lucia campus of the University of Queensland in Brisbane, Australia, showed that the NN ensemble was able to outperform modestly the SVR, obtaining solutions around 1% better after applying the filter feature selection method proposed.

In [Mohammadi2016] Mohammadi et al. applied an ANFIS system to select the most influential variables in a daily horizontal diffuse solar radiation prediction problem. Relevant variables were considered in order to study how different groups predict solar radiation: daily diffuse

and global solar radiation on a horizontal surface, sunshine duration, minimum air temperature, maximum air temperature, average air temperature, relative humidity, water vapor pressure, daily maximum possible sunshine duration, solar declination angle and extraterrestrial solar radiation on a horizontal surface. Four years of measured data from the Iranian Meteorological Organization (January 2009 to December 2012) were used. Analyses of the best combination of 3 variables were carried out, showing the percentage of improvement when considering 1, 2 or 3 variables. The best result obtained (3 variables) was over 30% better than the solution of the system considering just 1 variable. No comparison results with alternative approaches were shown in the work, which makes difficult a further analysis of the proposal's performance.

In [Will2011], Will et al. used a hybrid niching GA-LR approach to estimate global solar radiation in El Colmenar, Argentina. Data from 14 different weather stations were used in this work. Climatic variables such as daily average temperature, air humidity, atmospheric pressure, cloudiness, and sunshine hours were considered. The idea was to reconstruct the global solar radiation from the data in other 13 measurement stations. A niching GA with binary encoding was considered, where a 1 stands for including a given variable in the prediction, and a 0 stands for not including it. The LR was used due to its good computational performance and its interpretability, though the authors admit that NNs could obtain better results. The prediction obtained was only compared with the niching GA with different number of individuals and generations, with the best prediction result obtained with 200 individuals and 150 generations, improving a 3% the solution of the algorithm with 50 individuals and 35 generations. A complete comparison with alternative approaches was not provided.

In [Aybar2016], Aybar et al. applied a GGA to select the optimal set of features that maximizes the performance of an ELM for solar radiation prediction. As input variables, the system used the output of a numerical weather meso-scale model (WRF), i.e. variables predicted by the WRF at different nodes nearby the objective station. Experiments with real data from the Radiometric Observatory of Toledo (Spain) showed the good performance of this approach, in different prediction time-horizons, from hourly to 3 hours ahead prediction. Comparison with the results of an ELM showed that the feature selection procedure was able to improve the performance of the system over a 10% in terms of RMSE. In a related approach, Salcedo et al. in [Salcedo2017], proposed a novel CRO-SP algorithm for a FSP related to global solar energy prediction in Spain. In this case, the CRO with species was able to select the optimal number of input features for an ELM algorithm. This approach was compared against different alternative regressors: an ELM, a hybrid GA-ELM and also the GGA-ELM, in the global solar radiation Toledo's data, showing a significant improvement (over 21%) of performance versus those alternative approaches.

In [Garcia2018], García et al. evaluated three non standard multivariate feature selection approaches for a problem of solar energy prediction in Spain. Novel methods which automatically select the most adequate features were compared, adapting strong regression algorithms such as SVRs or DNNs for feature selection. Comparison with classical feature extraction approaches such as PCA or Lasso was carried out. Performance improvements over 5% were reported when

using the machine learning algorithms with feature selection versus the standard techniques.

Finally, note that Yadav et al., in [Yadav2014], offered a first general review of some works dealing with relevant parameters selection in solar energy prediction problems, in a larger framework of solar energy prediction with NNs.

2.1.3 Feature selection in energy-related problems

The last works discussed in this section are focussed on problems which are related to energy, but are not direct applications for predicting the main renewable sources. For example, in [Ahila2015], Ahila et al. classified power system disturbances with hybrid system including ELMs and PSO. In this case, the PSO approach was used to select the best features to serve as inputs of the classifier (ELM), and also the number of hidden nodes to enhance the performance of a multi-linear regression algorithm. Therefore, this system exploits a wrapper approach for classification disturbances in power systems. The experimental results showed that the proposed algorithm is faster and more accurate than alternative approaches in discriminating power system disturbances, producing improvements around 8% in classification accuracy. The database used included ten different power disturbances (10 classes classification problems), such as voltage sag, voltage swell, interruption, harmonics, flicker, oscillatory transient, sag with harmonics, swell with harmonics, notch and spike.

There are some recent works dealing with FSPs applied to electricity load forecasting. In [Koprinska2015] Koprinska et al. applied different filter feature selection methods to a problem of short-term electricity load from previous samples. The filter methods considered were based on autocorrelation of samples, Mutual Information, RReliefF and correlation-based techniques. The features selected were previously applied to different regressors such as NNs, model tree rules or linear regressors, in order to obtain accurate prediction of electricity loads. Two years of Australian electricity load data were used to show the goodness of the filter methods for feature selection, obtaining improvements between 3% and 4% depending on the method considered.

In [Jurado15], Jurado et al. applied different machine learning methodologies to forecast hourly energy consumption in buildings. In all cases, hybrid methodologies combining filters for feature selection (based on entropy measurements) with machine learning methods such as Fuzzy Inductive Reasoning, RF and NNs, were used. Experiments with actual data were carried out in Catalonia, Spain, where the different methods have been successfully compared with a traditional statistical technique (an ARIMA model for energy consumption prediction). The results showed that AI-based methodologies outperformed classical techniques, by far when the feature selection step is considered, with improvements over 20%.

Hu et al., in [Hu2015], solved a problem of mid-term electricity loads prediction by using a MSVR approach and a memetic algorithm for feature selection. This is a wrapper proposal, where the SVR forecasts the electricity load, whereas the memetic algorithm looks for the best set of features for the problem. A memetic algorithm is a hybrid approach composed by a global search algorithm and a local search procedure to enhance the search. In this case, the

global search algorithm was a firefly algorithm, and a local search based on a ranking of the best solutions was applied. The performance of this framework to predict daily interval electricity demands was tested in real data from North America and Australia, where the MSVR approach obtained important improvements (over 30%) over alternative algorithms based on NNs and classical SVR without feature selection mechanisms.

Finally, also related to the production of renewable energy in wind farms, Jiang et al. [Jiang2011] tackled a problem of feature extraction in wind turbine fault diagnosis (vibration signal analysis of wind turbines). A denoising method based on adaptive Morlet wavelet and Singular Value Decomposition was applied to feature extraction, producing variables which are much more interpretable than the previous ones.

2.1.4 Remarks on the FSP for energy related problems

The application of feature selection mechanisms improves the performance of machine learning prediction in renewable energy-related problems. The improvement fully depends on the specific application considered, but could be very important, in the range 5% to 40% of the prediction accuracy, just by including the feature selection algorithm. In all cases, feature selection implies an extra computational burden for the prediction systems. Filter methods, which use an external measure previous to the application of the prediction algorithm, are the less computationally demanding approaches. On the other hand, wrapper approaches may suppose a heavy burden for the algorithm, and they are usually combined with fast training prediction approaches. This can be seen in Table 2.1, which shows the most important articles applying feature selection and related methods to wind and solar energy problems. A first analysis of the table indicates that wrapper approaches are the most used, mainly in wind energy applications. Filter feature selection methods have had some impact in solar and energy-related applications. However, since they usually have less accuracy than wrapper approaches, they are also less used. The second interesting conclusion which can be drawn from this table is the large amount of algorithms used in wrapper FSP. Regarding prediction approaches, it is interesting that fast-trained approaches are more used, as previously suggested. In this sense, there are several works which bet for linear regression approaches, though in the last years, fast-trained neural networks like ELMs have been used more frequently.

Another important aspect to be discussed is related to the global search algorithms for wrapper feature selection. As can be seen in Table 2.1 there is a large amount of techniques applied, such as different types of GAs, PSO, DE, HS, CRO, greedy approaches or exhaustive search. The latter is only possible in those FSP where the number of features involved is small. In problems with a large amount of features, the search space size is huge (it grows at least as 2^N, where N is the number of features considered), so exhaustive search is not a computational viable option. This is why meta-heuristic approaches are mainly considered for wrapper feature selection systems. As Table 2.1 reveals, the number and characteristics of meta-heuristics approaches used in FSPs is large, and, moreover, it is difficult to decide which technique is better,

Table 2.1: Summary of the most important articles applying feature selection methods in energy applications.

Reference	FSP type	Predictor	FSP algorithm
Wind energy			
[Jursa2007]	wrapper	NN	PSO
[Jursa2008]	wrapper	NN, K-NN	PSO,DE
[Gupta2011]	wrapper	NN	GA
[Kou2014]	wrapper	GP	Greedy
[Salcedo2014b]	wrapper	ELM	CRO
[Salcedo2015c]	wrapper	ELM	CRO (HS)
[Carta2015]	filter	NN	Linear Correlation Coefficient
[Carta2015]	wrapper	NN	Exhaustive search
[Kong2015]	feature-extraction	SVR, RSVR	PCA
[Kumar2015]	filter	NARX	ReliefF
[Zhang2016]	wrapper	NN, SVR	LR
[Jiang2017]	filter	SVR	Kullback-Leibler divergence
[Zheng2017]	wrapper	ELM	PSO, GS
[Zhang2017]	wrapper	ELM	BSA
[Feng2017]	feature-extraction	NN,SVR,RF	PCA, recursive elimination
Solar energy			
[Fu2013]	feature-extraction	LR	Image Processing techniques
[Yadav2014b]	wrapper	NN	Exhaustive search
[Wang2015]	filter	SVM	correlation-based
[Rana2016]	filter	NNs, SVR	correlation-based
[Mohammadi2016]	Intrinsic	ANFIS	ANFIS
[Will2011]	wrapper	LR	niching GA
[Aybar2016]	wrapper	ELM	GGA
[Salcedo2017]	wrapper	ELM	CRO
[Garcia2018]	Filter	DNN, SVR	correlation-based

since there is not an unified framework for FSP, and each problem/application has different properties, so it fully depends on the specific type of FSP tackled.

2.2 Long-term wind speed variability

Long-term wind speed prediction and variability studies are very useful in different energy-related tasks, such as wind farm location and prospective works, maintenance planning, financial estimates or long-term electricity generation prediction [Ringkjob2017]. However this, long-term wind speed analysis has received little attention in the literature.

Several studies are focused on variability studies of monthly and seasonal wind speed. In [Ringkjob2017], ERA-Interim reanalysis data are used to model part of the long-term variability of the wind energy resource in central Europe by reconstructing surface wind distributions at different long-term scales. In [Robert2013] an adaptive general regression neural network is used to carry out spatial regression of monthly wind speed in the Alps considering NCEP/NCAR

reanalysis data. In [Tar2008] an analysis of the variability of monthly average wind speed in Hungary is carried out. The methodology uses a Weibull distribution to model the wind speed at different altitudes, showing that this distribution seems optimal at monthly scale. In [Kirchner2013] a study of multi-decadal variability of daily wind speed is carried out using NCEP/NCAR reanalysis data. In this case, an Evolutionary Computation algorithm was used to obtain different classes of wind speed, and study their inter-decadal variability. This work was extended in [Kirchner2015], where the long-term variability of wind power in Iberia was analyzed.

Other papers are focused on the statistical evaluation of monthly, seasonal or annual wind speed features in a given zone. For example, in [Ahmed2018] a statistical analysis on the monthly wind characteristics, using measured wind speed data for a five years period in the Southern Egyptian desert is carried out. The results show that a 150Mw wind farm could be successfully planned in the region to supply energy covering the necessities of the nearest cities. In [Soulouknga2018] an analysis of monthly average wind speed data in Chad is performed using Weibull distributions. In [Aquila2018] the authors try to identify whether statistically significant differences exist among the monthly wind average speed of the four major Brazilian states on wind energy productive capacity. They apply the Nested Gage Repeatability & Reproducibility, generally used on manufacturing quality management, to the wind speed series of the different States in order to identify major differences among them.

The analysis of the wind speed climatology in different regions of the World, including climate change, is the purpose of other studies. For example, in [Bianchi2017] an analysis of the South American long-term wind speed climatology is carried out, taking into account data from MERRA reanalysis and the effect of the main climate drivers in the zone, such as the AO or the ENSO. In [Pryor2006] long-term variability of wind indices is discussed. In [Cannon2015] a study of wind extremes for wind power generation in England is carried out with reanalysis data from the last 30 years. Studies of the climate change impact on wind energy always consider long-term analysis of wind speed data [Pryor2010, Pryor2005, François2017, Pryor2011].

2.3 A review on applications of the CRO algorithm

The CRO algorithm and its variants, described in Section 1.2.2, have been successfully applied to a large number of optimization and prediction problems. We review here the most important applications of the algorithm and its different versions, with special detail in problems related to the topics of this Thesis.

2.3.1 Applications of the CRO in energy-related problems

There are many different problems related to energy which can be stated as optimization problems. The first one we discuss here is the optimal layout of turbines in wind farms. This problem consists of choosing the best location of wind turbines in wind farms, in terms of

different optimization criteria and fulfilling a number of constraints [Serrano2014]. The CRO has been applied to a problem of turbines layout for of offshore wind farms (wind farms situated in the sea) in [Salcedo2014]. The basic CRO algorithm showed advantages against other meta-heuristics in this problem, specifically EA, DE and HS algorithms. The case-studies to test the CRO were real data from a site for offshore wind farm location in the North sea (Wikinger, Germany). Figure 2.1 (a) shows the Wikinger area where the wind farm is located, and the possible points to install the wind turbines. A design with 20 turbines was carried out, using as objective function the maximization of the Annual Energy Production (AEP) in the wind farm. Figure 2.1 (b) shows the layouts obtained with an evolutionary algorithm (AEP 84.256 Mwatt) and Figure (c) the result with the CRO (with an AEP of 84.352 Mwatt).

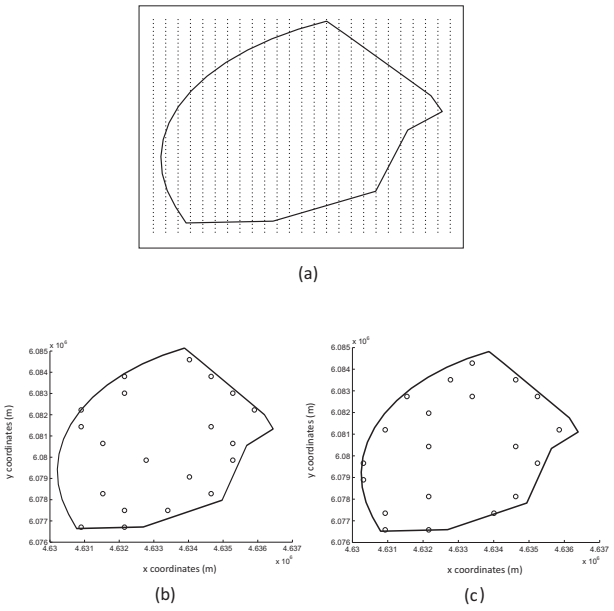


Figure 2.1: Wikinger wind farm (North sea); (a) Wind farm contour and possible location points; (b) Best solution obtained with an EA; (c) Best solution obtained with the CRO approach.

Wind speed prediction is another problem where the CRO has been successfully applied [Salcedo2015c]. In this case, the CRO has been used as part of a feature selection of the best variables to feed a neural network (an ELM, specifically) to carry out the prediction. In this case, the exploration of the algorithm is carried out by means of a HS search procedure as broadcast spawning, instead of using a classical crossover. The inclusion of this search pattern seemed to improve the performance of the CRO-ELM in this problem of wind speed prediction. In a similar approach, the CRO algorithm has been applied to solar radiation prediction in [Salcedo2014e]. In this case, the approach is slightly different from the application in wind speed prediction, since now the number of predictive variables (inputs) is small, and there is not a real need for feature selection. On the contrary, the CRO has been applied to slightly modify

the ELM weights, in such a way that the ELM prediction is improved. Results in real data from the radiometric observatory of Murcia (Spain) showed that the hybridization of the ELM with the CRO was able to improve the performance of the former in this problem of solar radiation prediction.

Table 2.2: Variables considered in the problem of energy demand estimation at a nation level.

#	variable
1	GDP
2	Population
3	Export
4	Import
5	Energy Production (kTOE)
6	Electricity power transport (kWh)
7	Electricity production(kWh)
8	GDP per unit of energy use
9	Energy imports net (% of use)
10	Fossil fuel consumption (% of total)
11	Electric power consumption (kWh)
12	CO_2 emissions total (Mtons)
13	Unemployment rate
14	Diesel consumption in road (kTOE)

Another application of the CRO approach in energy is focused on a problem of total energy demand estimation at nation's level [Salcedo2017b]. The problem consists of estimating (within a given time-horizon prediction) the energy that will be consumed in a country using macro-economic variables. This problem had been previously tackled using different meta-heuristics [Ceylan2004, Kiran2012, Salcedo2015], but the application of the CRO brought novelties, both in the problem's resolution, and also in terms of meta-heuristics design. In [Salcedo2017b] the concepts of CRO with species and substrates were introduced to tackle this problem of energy demand estimation, concepts that have been developed later on to obtain new co-evolution approaches. This problem can be stated as estimating the total energy demand in a country (Spain) with a time horizon of 1 year: $f = E^{k+1}(\mathbf{x_k})$, where the predictive variables for time k are matched with the energy demand for time $k+1$. Data from 1980 to 2011 are available, with a total of $m = 14$ predictive variables described in Table 2.2. A maximum number of features is set ($m' = 4$ in this case) for all models considered, so note that a feature selection mechanism is applied by using a binary encoding. A set of $T = 6$ different models for the calculation of the energy demand estimation were taken into account, and the parameters of each model (w_x) are also considered in the corals' encoding:

- 1. Linear model (*lin*):

$$s_{1,\mathbf{w}} = w_1 + w_2 \cdot x_1 + w_3 \cdot x_2 + w_4 \cdot x_3 + w_5 \cdot x_4 \tag{2.1}$$

- 2. Exponential model (*exp*):

$$s_{2,\mathbf{w}} = w_1 + w_2 \cdot x_1^{w_3} + w_4 \cdot x_2^{w_5} + w_6 \cdot x_3^{w_7} + w_8 \cdot x_4^{w_9} \tag{2.2}$$

- 3. Logarithmic model (*log*):

$$s_{3,\mathbf{w}} = e^{(w_1 + w_2 \cdot |ln(x_1)|^{w_3} + w_4 \cdot |ln(x_2)|^{w_5} + w_6 \cdot |ln(x_3)|^{w_7} + w_8 \cdot |ln(x_4)|^{w_9})} \tag{2.3}$$

- 4. Quadratic model, version A (*qua*):

$$s_{4,\mathbf{w}} = w_1 + w_2 \cdot x_1 + w_3 \cdot x_2 + w_4 \cdot x_3 + w_5 \cdot x_4 + w_6 \cdot x_1 \cdot x_2 + w_7 \cdot x_1 \cdot x_3 +$$

$$+ w_8 \cdot x_1 \cdot x_4 + w_9 \cdot x_2 \cdot x_3 + w_{10} \cdot x_2 \cdot x_4 + w_{11} \cdot x_3 \cdot x_4 + w_{12} \cdot x_1^2 + w_{13} \cdot x_2^2 + w_{14} \cdot x_3^2 + w_{15} \cdot x_4^2 \tag{2.4}$$

- 5. Quadratic model, version B (*qub*):

$$s_{5,\mathbf{w}} = w_1 + w_2 \cdot x_1^{w_3} + w_4 \cdot x_2^{w_5} + w_6 \cdot x_3^{w_7} + w_8 \cdot x_4^{w_9} + w_{10} \cdot x_1 \cdot x_2 +$$

$$+ w_{11} \cdot x_1 \cdot x_3 + w_{12} \cdot x_1 \cdot x_4 + w_{13} \cdot x_2 \cdot x_3 + w_{14} \cdot x_2 \cdot x_4 + w_{15} \cdot x_3 \cdot x_4 \tag{2.5}$$

- 6. Mix model (*mix*):

$$s_{6,\mathbf{w}} = w_1 + w_2 \cdot e^{(w_3 + w_4 \cdot x_1 + w_5 \cdot x_2 + w_6 \cdot x_3 + w_7 \cdot x_4)}. \tag{2.6}$$

With these prediction models, an error function is needed to measure the quality of the prediction. For example the Mean Square Error, defined, for a given model h as:

$$\epsilon_{MSE} = \frac{1}{N} \sum_{j=1}^{N} \left(E^{k+1}(\mathbf{x_k}) - s_{h,\mathbf{w}} \right)^2, \tag{2.7}$$

Note that the use of different models complicates the problem, since each model implies a different encoding, with different length (for example model $s_{1,\mathbf{w}}$ induces an encoding of length 19 (14 binary variables plus 5 from $\mathbf{w}$), and model $s_{5,\mathbf{w}}$ encoding length is 29 (14 binary + 15 from $\mathbf{w}$)). In this case, the use of the CRO-SP described in Section 1.2.3 and CRO-SL (Section 1.2.3) can solve the problem. In the CRO-SP each species stands for a given model s, and the encoding of the corals in the reef is the following:

$$\Xi_{i,j}^t = [t|\mathbf{I}|\mathbf{w}], \tag{2.8}$$

where t stands for the species number (model to calculate the energy estimation), $\mathbf{I}$ is the binary part for feature selection and $\mathbf{w}$ stands for the model's parameters. On the other hand, in the CRO-SL, the encoding is as follows:

$$\Xi_{i,j} = [\mathbf{I}|\mathbf{w}]\,, \tag{2.9}$$

where $\mathbf{I}$ is the binary part for feature selection, and $\mathbf{w}$ stands for the parameter of the models. The length of $\mathbf{w}$ is equal to the maximum length required by the largest model considered (all corals in the reef have the same encoding, and the substrate layer stands for the different models considered in the problem, i.e. now each substrate represents a different model, and each coral is evaluated with a different objective function to estimate the energy demand depending on the substrate it falls).

As previously mentioned, in [Salcedo2017b] the specific case of energy demand estimation in Spain from macro-economic variables has been tackled. The CRO-SP and CRO-SL have been used, obtaining good results, slightly better in the case of the CRO-SL as can be seen in [Salcedo2017b]. From the algorithms' point of view, the convergence of both approaches is different. Figure 2.2 shows the differences in convergence of the CRO-SP and CRO-SL in this problem. As can be seen, in the CRO-SP (Figure 2.2 (a)), the worse species perish with the algorithm's generations, and finally only the best one (the best energy model in this case) remains in the reef. The effect of fitness increasing is because the best individual of a given species is killed by another individual (from a different species). This can also be seen if we represent the number of corals belonging to each species in the reef (Figure 2.3). As can be seen, the number of corals belonging to each species is variable in the reef evolution. On the other hand, Figure 2.2 (b) shows the evolution of the CRO-SL, and as can be seen, in this case all the substrates remain equal until the end of the evolution, so the effect of species extinction is not present here. It is important to note that both versions of the CRO algorithm select the same model for energy demand estimation (the *qua*, given by Equation 2.4).

Finally, in [Salcedo2016] the CRO-SL algorithm has been successfully applied to a problem of battery scheduling in a micro-grid with renewable generation (wind and photovoltaic generations) and variable prices of electricity. The CRO-SL approach was compared with a *deterministic* use of the battery, in which it is charged with the maximum possible power every period of time in which the generation is larger than the load's demand. On the contrary, in the periods of time in which the load's demand is larger than the generation, the battery is discharged with the maximum possible power. Five substrates were defined in this application for the CRO-SL: HS, DE, two-points crossover (2Px), multi-point crossover (MPx) and Gaussian mutation (GM). The CRO-SL approach was able to clearly improve the deterministic use of the battery in all the tested cases (different weeks periods in different seasons of the year). Figure 2.4 shows the results obtained by the CRO-SL in comparison with those by the deterministic use of the battery, for a winter week in the micro-grid considered. Note that the use of the CRO-SL for battery scheduling instead of its deterministic use produces an important reduction

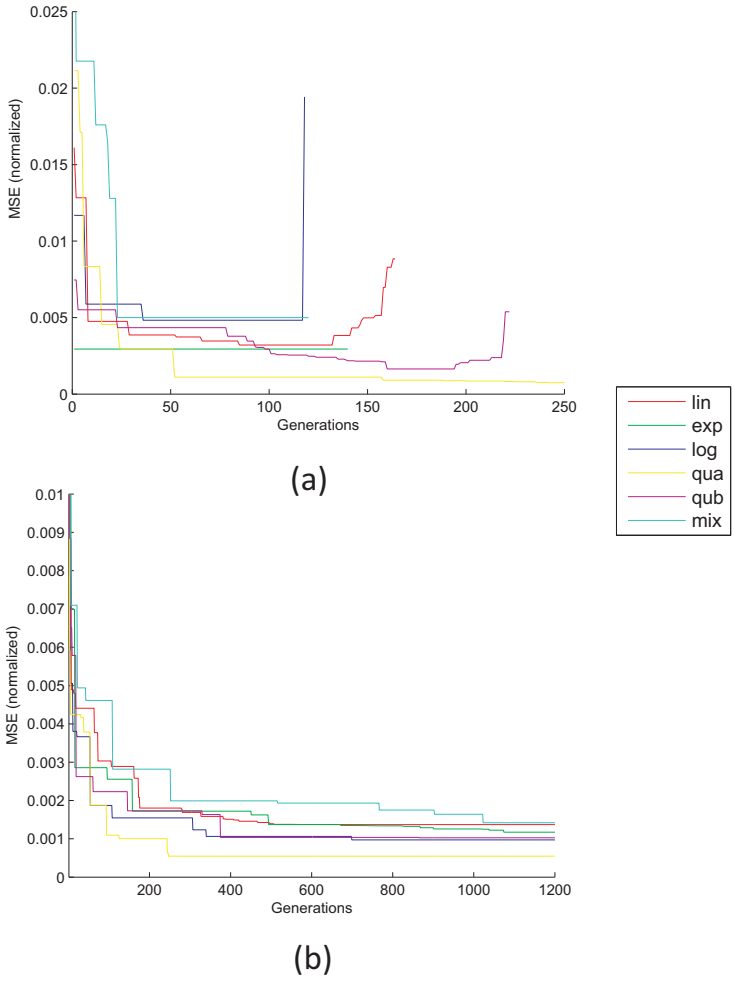


Figure 2.2: Evolution comparison between the CRO-SP and CRO-SL in the problem of energy prediction from macro-economic variables in Spain; (a) CRO-SP; (b) CRO-SL.

of the electricity consumption from the main grid, specially in time instants where the electricity price is high. Figure 2.5 shows the relative importance of each substrate in the CRO-SL in this problem, in terms of the percentage of times in which a given substrate provides the best larva (new solution) in each generation. It seems that the crossover operators are useful in this application, and also the DE substrate contributes to the evolution, whereas the GM provides little contribution and the HS substrate does not provide a good exploration of the search space in this case.

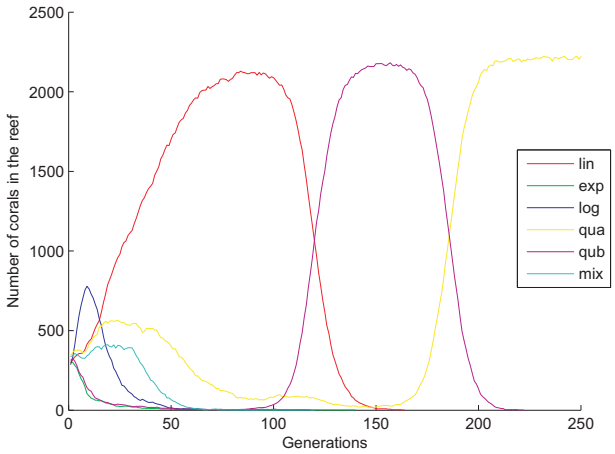


Figure 2.3: Evolution of the number of corals within generations (CRO-SP) in the problem of energy prediction from macro-economic variables in Spain.

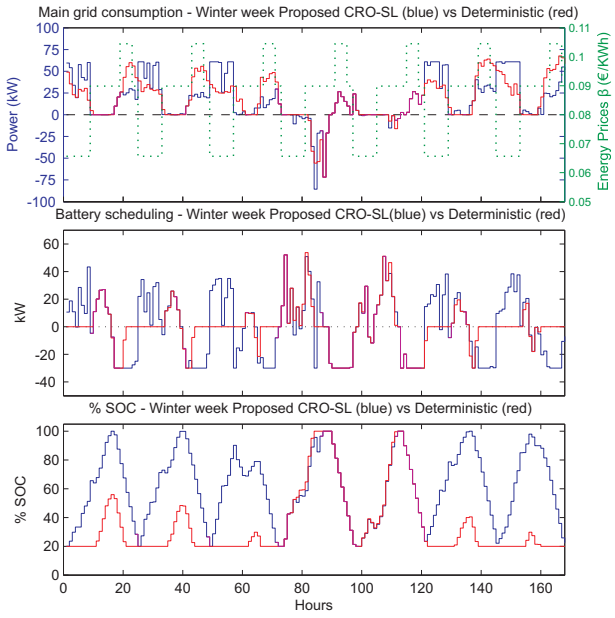


Figure 2.4: Comparison of the results obtained using the proposed CRO-SL (blue line) and the deterministic approach (red line) for a winter week in the micro-grid considered; (a) Consumption from the main grid, (b) Battery power scheduling and (c)% SOC (State of Charge) in the battery.

2.3.2 Alternative applications of the CRO algorithm

Recently, alternative applications of the CRO algorithms have been successfully introduced in different research areas. For example, different versions of the CRO have been applied to

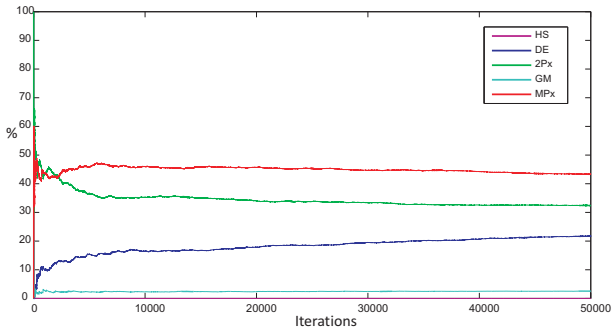


Figure 2.5: Comparison of the effect of the different substrates of the CRO-SL in the Winter week considered (percentage of times in which a given substrate gives the best larva (new solution) in each generation).

Telecommunication related problems. In [Salcedo2014f] the CRO is introduced to tackle the problem of optimal distribution of different services over available technologies in mobile communications systems, in order to obtain an optimized-cost network deployment. Different services within all the spectrum of mobile communications technologies (2G-GSM, 3G-UMTS, 3G-HSPA and 4G-LTE) have been considered. The CRO obtained improvements in terms of the network deployment cost, improving the results of a classical evolutionary algorithm and a Teaching Based Learning meta-heuristic approach. In [Salcedo2014g] another problem of optimal network deployment for 2G-GSM systems is tackled with the CRO algorithm with grouping encoding. In this case, the optimal location of Base Stations (BSs) is carried out, in terms of three design objectives: maximization of the network coverage, minimization of the installation cost, and minimization of the electromagnetic pollution caused by the installation of new BSs. A CRO with a specific grouping encoding [Falkenauer1992] was applied, obtaining a solid algorithm which is able to manage several solutions involving a different number of BSs in the same population. The performance of the grouping CRO in this problem was tested in comparison to alternative grouping versions of EAs, HS and PSO algorithms. In [Salcedo2016b], a different version of this problem was also tackled with the grouping CRO approach. In this case, a capacity constraint in BSs was considered, in such a way that the number of users in each BS is restricted. The CRO obtained again very good performance. In [Li2016] the coverage optimization problem in a directional sensor network was formulated as a multi-objective optimization problem. The problem's definition takes into account the coverage rate of the network, the number of working sensor nodes and the network's connectivity. A Tchebycheff decomposition method is introduced in this paper making possible to decompose the multi-objective problem into a single-objective problem. A novel LA-CRO is then introduced to solve the problem, in such a way that the LA-CRO obtains the best set of parameters for this problem. The LA-CRO approach was tested in several computational tests on this problems, showing a good performance. In [Ficco2016] a cloud resource allocation problem with the CRO is tackled. In this case, the CRO is used

to model cloud elasticity in a cloud-data center and on the classic Game Theory to optimize the resource reallocation schema with respect to cloud provider's optimization objectives and customer requirements.

There are other research areas where the CRO and its versions have obtained good results. For example, in [Yang2016] a hybrid DECRO has been applied for the optimal training of ELM networks. Experimental results in different sets have shown that DECRO-ELM can reduce the prediction time of original ELM, and obtain better performance for training ELM than both DE and CRO on their own. In [Medeiros2015] the CRO performance in clustering-related problems has been studied. The CRO has been adjusted to provide a good clustering partition for different datasets. Three new modifications of the CRO algorithm and an index to be used as objective function for the problems have also been introduced in [Medeiros2015]. A comparison of the different CRO version with a hybrid genetic algorithm for clustering is carried out to show the performance of the CRO in these problems. In [Silva2016] the CRO is used to combine multiple partitions generated by different clustering algorithms into a single clustering solution. The CRO was able to improve the solution of a classic GA for this task. In [Pichpibul12] a novel version of the CRO for the capacitated vehicle routing problem is introduced. The modification of the CRO algorithm consists of using an initial solution for the problem based on the probabilistic Clarke-Wright savings algorithm [Pichpibul12b], and then improve the obtained solution with the CRO mechanism. This approach has been tested in 14 well-known capacitated vehicle routing benchmark instances, and compared it with other existing algorithms in the literature, obtaining competitive results. In [Deniz2016] a discussion on the CRO suitability for the dynamic cell formation problem is carried out. In [Yawei2016] an application of the CRO to the optimal parameters' identification of a permanent magnet synchronous motor is presented. The CRO has been compared against a PSO algorithm with least squares, outperforming this approach in obtaining the best set of parameters for the motor. Recently, two works have described the performance of the CRO and CRO-SL in a problem of medical image registration [Bermejo2017, Bermejo2018]. In this case, the CRO or CRO-SL is used to estimate the optimal transformation of a medical image to fit to an objective image. In both cases, the CRO-based algorithms have been benchmarked with state-of-the-art evolutionary and non-evolutionary image registration methods, showing their good performance in all cases. The CRO-SL algorithm has been applied to different problems of vibration cancellation in buildings. Specifically, in [Salcedo2017] the CRO-SL has been applied to a problem of Tuned Mass Dumper (TMD) design, a passive vibration cancellation method used in many buildings. In this case, the CRO-SL is used to adjust the parameters of the TMD to cancel the required frequencies in a building. The evaluation of the CRO-SL performance has been carried out with simulation and also in a constructed laboratory model. In [Camacho2018] the CRO-SL has been applied to the design of an active method for vibration cancellation. In this case, the CRO-SL is used to tune an Active Vibration Control (AVC), via inertial-mass actuators. It is shown that the AVC optimization with the CRO-SL is a viable technique to mitigate human-induced vibrations in civil structures. Finally, the basic version of the CRO has been applied to problems of time

series segmentation in [Duran2017, Duran2018]. In [Duran2017], the CRO has been applied to the segmentation of specific time series of wave height, for marine energy applications. In [Duran2018], the study is deeper, and it includes the optimal segmentation of different times series with a statistically-driven CRO version, where some of the CRO parameters are set by means of statistics of the reef solutions.

Chapter 3

Feature section in wind energy prediction systems

3.1 Introduction

Wind energy is currently the most important sustainable energy source in the World, in terms of annual growing, penetration in the power system and economic impact [Kumar2016]. One of the problems of wind energy is that it exhibits intermittent generation (depending on the weather conditions) [Yan2015], which makes difficult its integration in the system. Wind power generation forecasting is therefore a key factor to improve this integration [Tasci2014, Salcedo2015c, Renani2016, Capellaro2016, Munteanu2016].

Nowadays, modern forecasting models for short-term power prediction (or, equivalently, the wind speed) in wind farms are based on the combination of physical and statistical models [Costa2008]. The physical models can be global, meso-scale or even local, taking into account the specific orography of the wind farm [Landberg1999, Landberg2001]. Statistical models are usually included in prediction systems jointly with physical models. It has been shown that they produce a significant improvement in the prediction when compared with purely physical approaches. In the last few years, many different statistical approaches have been applied to wind speed prediction, including linear prediction models [Riahy2008], classical Box-Jenkins methodologies such as auto-regressive models [Torres2005] and other time series analysis such as the Mycielski algorithm [Hocaoglu2009], different clustering algorithms [Kusiak2010b], and several modern computational approaches such as neural networks [Barbounis2007, Li2001, Bilgili2007, Salcedo2009, Salcedo2009b, Li2010], neural networks ensembles [Kusiak2010], Bayesian methods [Jiang2013], Support Vector Machines [Mohandes2004, Salcedo2011, Ortiz2011], or hybrid methods, combination of the previously mentioned techniques: neural networks and auto-regressive models [Khashei2009], auto-regressive models and Kalman filtering [Liu2012], neural networks and Markov models [Pourmousavi2011], wavelets and neural approaches [Zhang2013, Liu2013], etc.

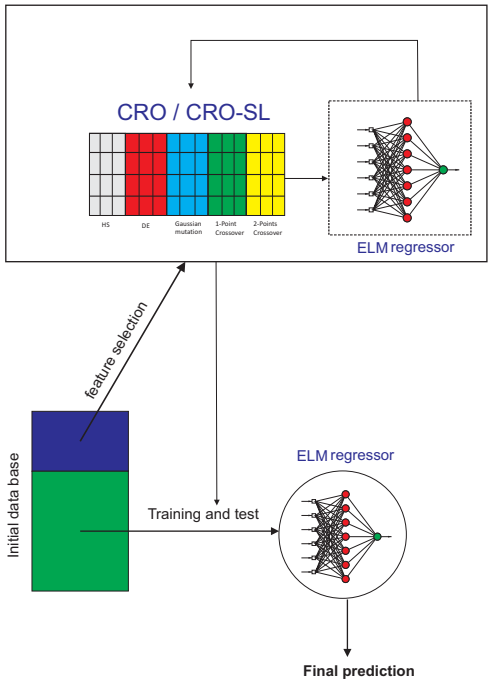


Figure 3.1: Hybrid CRO-SL-ELM system for wind speed prediction with feature selection (see text for details).

Many of the statistical approaches included in short-term wind speed prediction systems use the output of physical models (meteorological variables) as input, trying to improve the quality of the prediction result. Meteorological variables measured or modeled at different points can be used as inputs for the statistical methods, such as wind force and direction, temperature, pressure, etc. In fact, even if we consider a reduced-size grid, the number of available meteorological variables (from a given physical model), is huge, and some variable selection is needed, as pointed out in Section 2.1. We therefore deal with a problem of feature selection in a wind speed prediction system, based on data from numerical models, with an statistical final approach given by a prediction algorithm. The two systems applied follow the structure given in Figure 3.1, and consist of a hybrid approach, mixing a meta-heuristic algorithm for feature selection (CRO or CRO-SL) with an ELM for prediction. They are, therefore, wrapper approaches for the FSP. After the FSP, the wind speed prediction is obtained by means of another ELM or alternative prediction mechanism, trained only with the selected variables after the FSP. Two different cases of study are included in this section: first, we evaluate the performance of the CRO approach in a FSP problem in a wind farm located at the west coast of the USA. Second, an enhanced CRO-SL algorithm is described for a FSP problem in Spain. Comparative results with alternative meta-heuristics and final prediction are provided in the discussion of these applications.

3.2 FSP with the CRO-ELM algorithm

We consider a grid Ω formed by $N \times N$ nodes and a given measuring tower $\mathfrak{M}$. We also consider a time series of wind speed values in $\mathfrak{M}$, and a times series of M meteorological variables (features) in each node of the grid, obtained from a given physics-based prediction model. The wind speed series in $\mathfrak{M}$, and the meteorological series in the points of the grid are synchronized in time. Note that for large values of M, the number of available meteorological variables is huge (may be over 5000). The problem consists of predicting the wind speed in $\mathfrak{M}$ by using the predictive meteorological variables of the grid points. Figure 3.2 shows an example of a grid Ω and measuring tower $\mathfrak{M}$. We consider then a fix number m of final meteorological variables (out of the total $n = M \times (N \times N)$) to do the wind speed prediction. In this case we have carried out experiments with different number of fixed variables m, so the objective of the problem is to obtain the best set of m variables that provides the best performance of the system in terms of wind speed prediction.

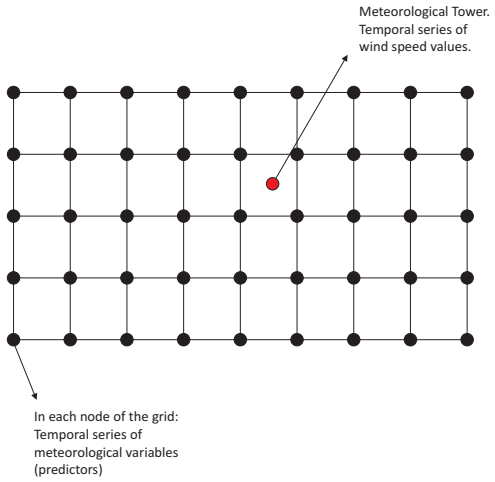


Figure 3.2: Example of the grid and measuring tower in a wind farm.

With this in mind, the encoding of each coral Ξ (problem's solution) in the CRO is the following: each meteorological variable included in the prediction system needs a total of four parameters to be identified, (x, y, id, ma), where x stands for the x-coordinate in the grid, y stands for the y-coordinate in the grid, id stands for the variable identifier and ma is a binary variable in such a way that a 1 means that we consider a moving average of the series of that variable, and a 0 means that it is not considered. The final encoding of a coral in the algorithm is therefore a $(m \times 4)$-length vector:

$$\Xi = [x_1, y_1, id_1, ma_1, \ldots, x_m, y_m, id_m, ma_m] \tag{3.1}$$

In order to test the performance of the CRO-ELM algorithm for short-term wind speed

prediction, we have carried out a number of experiments with real wind speed data from a measuring tower ($\mathfrak{M}$) in a wind farm in USA (see Figure 3.3). In the following sections we describe in detail the data used to evaluate the CRO-ELM performance, the predictive variables considered in this case, as well as a brief description of alternative algorithms we have used to contextualize the CRO-ELM analysis.

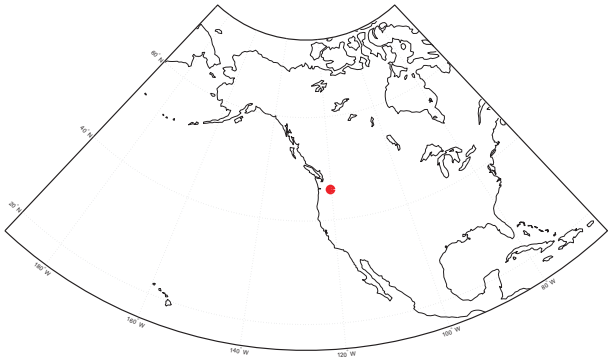


Figure 3.3: Location of the wind farm considered to test the CRO-ELM algorithm.

3.2.1 Data used, variables considered and methodology

One year of hourly 10m wind speed data (01/03/2007-29/02/2008) is considered. An 11×11 grid with a 5 km resolution surrounding the measuring tower is taken into account, and in each node of the grid, a series of 27 meteorological variables (at different height levels) is considered. Table 3.1 shows the predictive meteorological variables taken into account. Variables in the analysis period (01/03/2007-29/02/2008) have been obtained with a meso-scale WRF model [Skamarock2005] in backcast or hindcast mode, using the National Center for Environmental Prediction and National Center of Atmospheric Research (NCEP/NCAR) global Reanalysis dataset. The WRF is a powerful meso-scale numerical model prediction system designed for atmospheric research and also for operational forecasting needs. It was developed in collaboration by the NCAR, NCEP, FSL, AFWA, the Naval Research Laboratory, the University of Oklahoma, and the FAA of the USA. Note that there are 27 predictive variables, some of them are direct outputs of the meso-scale model and some others corresponding to derived variables (functions such as logarithms or exponential of direct variables). As mentioned, we also consider the possibility of a moving average on the variable series, so we finally have $M = 54$ possible variables to be selected in each grid point. Thus, the total number of variables involved in the problem is $n = M \times (N \times N)) = 54 \cdot 121 = 6534$ variables. We have first split the available data into a training (75% of the data) and a test set (25%). After this first splitting of the data, the following methodology has been carried out. The RMSE will be used as coral health (objective) function in the CRO algorithm. However, this RMSE must be calculated in the CRO considering

only the training data. In order to obtain an objective measure of RMSE which provides then the best generalization of the algorithm, we have included a procedure of n cross-validation, in which the training data is split again into n sets, and the ELM is trained with $n-1$ of these sets and tested in the remaining one, in such a way that all the sets are used as test set. The RMSE provided as health function of the coral will be the average of all the RMSE obtained for each of the sets. In a final step, after the best coral has been obtained in the CRO, we can obtain its final associated RMSE in the test set as the final result of the CRO-ELM. Note that all the results shown in this paper are referred to this final RMSE in the test set.

Table 3.1: Predictive meteorological variables used in the short-term wind speed prediction problem considered.

id #	Meteorological variable
direct measures	
0-5	wind speed and direction at different heights (0m, 10m, 20m, 50m, 80m, 100m)
6-10	wind direction at different heights (0m, 10m, 20m, 50m, 80m)
11-13	temperature at different heights (0m, 2m, 20m)
14	specific humidity (2m)
15	Sea Level Pressure
16	long wave down radiation (0m)
17	short wave down radiation (0m)
18	precipitation
19-26	function combinations (log, exp) of variables 0-18

3.2.2 Algorithms for comparison

We can establish two levels of comparison with the CRO-ELM. First, we can evaluate the goodness of the CRO as global searcher. In order to do it, we have used an EA [Eiben2003] to solve the same FSP associated with short-term wind speed prediction. In this case, the structure of the algorithm for comparison is exactly the same that the proposed approach, substituting the CRO by an EA. A standard EA with two-points crossover and Gaussian integer mutation has been used in this comparison. The second level of comparison is the evaluation of the ELM as a regressor. In this case, note that the computation time of any candidate to substitute the ELM in the approach must be extremely fast. Otherwise, the computation time of the algorithm could be completely unacceptable. Considering this important constraint, we have tested a LR to substitute the ELM in this second comparison carried out, to form a CRO-LR algorithm.

3.2.3 CRO-ELM results

Table 3.2 shows the results obtained (in terms of RMSE) by the CRO-ELM and EA-ELM approaches. It is easy to see that the CRO-ELM algorithm obtains better results than the EA-ELM in all the experiments carried out, with a small error in wind speed prediction. The average improvement over the EA-ELM algorithm is about 2%. This may seem a small figure, but it represents an important improvement in terms of energy production and obtained revenue for an average wind farm, as we will show in the discussion of this section. Note that the wind speed prediction system is able to obtain accurate wind speed prediction, with a mean square error in the test set around 2.5 m/s. It is interesting that the best result obtained with the CRO-ELM contains $m = 9$ variables. This means that we have reduced the total 6534 initial possible variables to just 9, keeping the accuracy in the prediction. This implies an improvement in the wind speed prediction system in terms of computational complexity, since once the system is trained, we will have to consider just 9 variables to obtain a good wind speed prediction.

Table 3.2: Results obtained with the CRO-ELM and EA-ELM in the FSP problem associated with short-term wind speed prediction.

# predictive variables	RMSE CRO-ELM [m/s]	RMSE EA-ELM [m/s]
4	2.57	2.61
5	2.56	2.59
6	2.55	2.59
7	2.53	2.58
8	2.54	2.57
9	2.50	2.56
10	2.50	2.57
11	2.51	2.58
12	2.52	2.58
13	2.51	2.59
14	2.51	2.59
15	2.53	2.60

In the second comparison carried out to evaluate the performance of the algorithm, we analyze a comparison between the CRO-ELM and CRO-LR approaches. In this case we focus on the case of 9 final variables. The result obtained by the CRO-LR is in this case 2.99 m/s

in RMSE. This results is poorer than the one obtained with the ELM, which seems to be more accurate with a similar computational cost. Figure 3.4 shows the best CRO evolution obtained, considering cross-validation RMSE with the ELM and LR regressors. We can also have a visual shot of the CRO-ELM performance in this problem by comparing the best wind speed prediction obtained against the real wind speed. Figure 3.5 shows such a comparison. It is easy to see that the CRO-ELM obtains an accurate reconstruction of the wind speed, missing some ramps, but following quite well the wind speed trend in general.

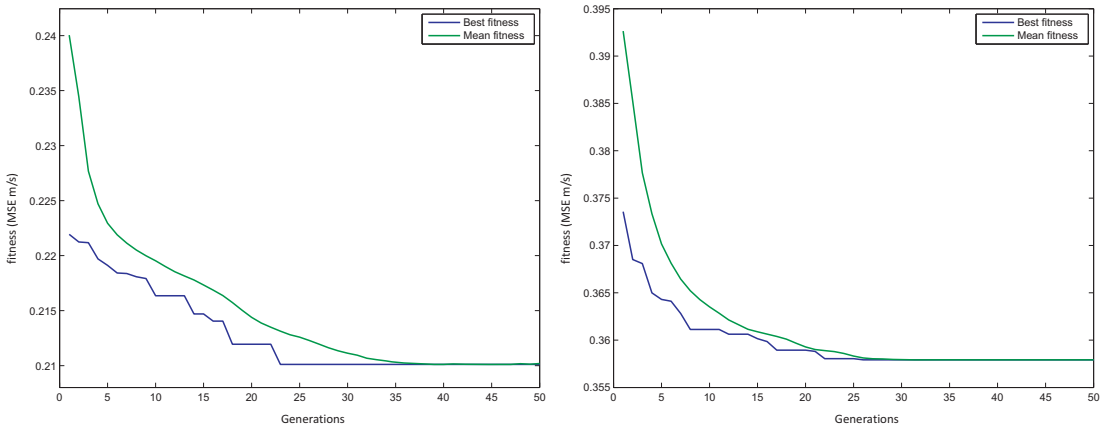


Figure 3.4: Best CRO-ELM and CRO-LR evolution (see text for details); (a) CRO-ELM; (b) CRO-LR.

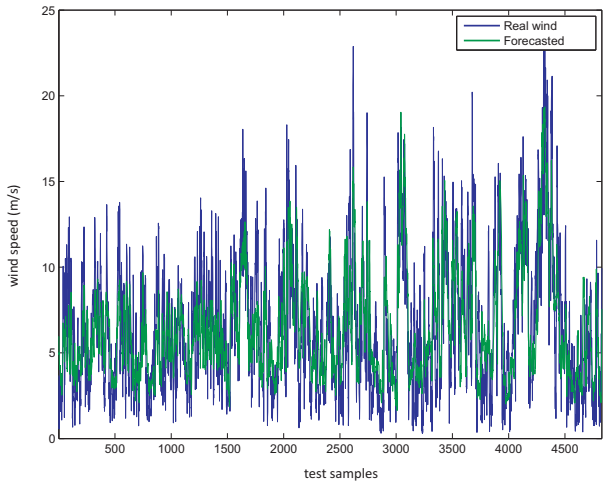


Figure 3.5: Wind speed prediction obtained with the CRO-ELM algorithm and real wind speed.

3.2.4 Further analysis on the CRO-ELM approach and discussion

In the previous section we have analyzed the performance of the ELM and LR evolution (using a CRO and EA algorithms), in order to fix a small number of predictive variables in a problem of wind speed forecast. As has been mentioned, the election of the ELM or LR algorithms is useful because the wrapper feature selection requires low computational approaches to be hybridized with the global search algorithms. There are, however, powerful algorithms for regression that could produce excellent results in wind speed prediction. One of these approaches is the SVR algorithm [Smola2004], that is known to be one of the most accurate existing regressor approaches, and it has been successfully applied to wind speed prediction previously [Ortiz2011, Salcedo2011]. On the other hand, this algorithm involves a high computation time, even harder if SVR parameters are sought previously to its application to the problem (in this case we obtain the SVR parameters by using a grid search approach). Thus, it is not directly applicable in hybridization with the CRO, but, we could, of course, test its performance by applying it to the features selected by the CRO-ELM or CRO-LR algorithms. The idea is to select the best reduced set of predictive feature with the CRO-ELM or CRO-LR, and then using the SVR to test this reduced set of features. We have carried out this experiment, in the case of a set of 9 features. Table 3.3 shows these results. Note that the SVR approach applied to the features selected either by the CRO-ELM or CRO-LR provides worse results than the ELM.

Table 3.3: Results obtained (RMSE in m/s) with the ELM and SVR approaches as prediction algorithms, using the features selected by the CRO-ELM and CRO-LR algorithms.

	ELM	SVR
CRO-ELM	2.50	2.92
CRO-LR	2.68	2.99

In order to analyze the impact that the improvement in wind speed prediction has in terms of economical revenues: we have obtained the energy production series from the wind speed series, by using the model in [Norgaard2004]. This model considers the behavior of complete wind farms and not only isolated wind turbines. We have assumed an average wind farm with 50Mw, with standard losses and wind turbines availabilities. Due to the non-linearity of energy production with the wind, in this case, the RMSE in energy is similar to the one obtained for the wind (about 2%). Considering the Spanish case, in which penalties for errors in wind energy production are different depending on the hour and there are differences in penalties for under and over estimation of production, an improvement of 2% in the wind speed prediction implies a revenue in the final wind energy price about 0.1% approximately. In Spain, for a standard 50Mw wind farm and a medium capacity of 26.5% this implies an extra income of 6500 Euros/year.

3.3 FSP with the CRO-SL-ELM algorithm

For the analysis of the CRO-SL-ELM, we consider real wind data from a wind farm located in northern Spain, Figure 3.6. Ten years of data are available, from 11/01/2002 to 29/10/2012, 80% of them will be used to train the prediction system, whereas the remainder 20% will be used for test purposes. Regarding predictive variables, a total of 98 meteorological variables serve as initial input parameters in this case. All the predictive variables have been obtained from the WRF in a single node, which corresponds to the wind farm situation. The CRO-SL-ELM is then applied as a system to obtain a wind speed prediction from these variables. Table 3.4 summarizes the complete set of variables from the WRF meso-scale model considered. Among the variables considered, we have chosen those which may have direct relationship with the wind speed in the objective site, for example wind speed and direction in the WRF nodes at different levels, pressure, temperatures at different levels, etc. We also consider some indirect measurements which are displayed in the table as $g(x)$, where g stands for a function such as logarithmic or exponential, and x stands for the corresponding direct predictive variable. In this case, the encoding of each coral Ξ is simpler than in the previous problem, since all the variables have been previously numbered. Then, a binary encoding is possible for Ξ in this case, where a 1 stands for considering the feature and a 0 for discarding it.

Figure 3.6: Wind farm considered for the experiments in the CRO-SL-ELM case.

The ELM is trained with an hourly prediction time horizon, and then a daily prediction from the average of hourly results is also obtained. Thus, we use different objective functions to evaluate the features selected by the CRO-SL:

$$f_1(\mathbf{x}) = r_h^2, \tag{3.2}$$

This first objective function only considers the Pearson coefficient r_h^2 for the wind speed hourly prediction.

$$f_2(\mathbf{x}) = 0.4 \cdot r_h^2 + 0.6 \cdot r_d^2, \tag{3.3}$$

The second objective function considers both the hourly and daily prediction time-horizons, using some weights to combine both predictions. The last objective function does the same, but with a small weight in the wind speed hourly component:

$$f_3(\mathbf{x}) = 0.1 \cdot r_h^2 + 0.9 \cdot r_d^2, \tag{3.4}$$

3.3.1 CRO-SL-ELM results

As initial reference values we show in Table 3.5 the results obtained by an ELM and a MLR algorithm [Brown2009] (used as baseline approach), with the 98 input variables, i.e. without any feature selection mechanism. As can be seen, the hourly prediction of the wind speed with these methods from numerical weather variables is good, obtaining correlation coefficient of 0.69 and 0.68, respectively. Table 3.6 shows the performance of the hybrid CRO-SL-ELM algorithm. A comparison with a CRO-ELM approach similar to that in [Salcedo2014b] but considering a binary encoding, and with a CRO-MLR and CRO-SL-MLR is also shown in this table. It is possible to see that the CRO-SL-ELM is able to improve the accuracy of the ELM on its own. This indicates that a process of feature selection is positive for improving the prediction capability of the system. Regarding the comparison with the CRO, Table 3.6 shows that the system with the CRO-SL obtains better results than the system with the basic CRO. It is an indication that the CRO-SL is able to better explore the search space. The results obtained also show that the hybrid CRO-SL-ELM outperforms CRO-MLR and CRO-SL-MLR, i.e. the ELM works in general better than the MLR as final prediction mechanism. The feature selection process is also positive when the MLR was used as final regression, improving the performance of the MLR on its own. Note that the solution provided by the basic CRO included 52 features out of the initial 98, but the CRO-SL provided a better solution with 25 features. Table 3.7 shows the best set of features according to the CRO-SL with f_1 objective function.

Table 3.4: Complete set of predictive features from the WRF model for wind speed prediction.

#	Variable	name
1	wind speed(10m)	ϑ_{10}
2	wind direction(10m)	θ_{10}
3	wind speed(20m)	ϑ_{20}
4	wind direction(20m)	θ_{20}
5	wind speed(50m)	ϑ_{50}
6	wind direction(50m)	θ_{50}
7	wind speed(100m)	ϑ_{100}
8	wind direction(100m)	θ_{100}
9	wind speed(200m)	ϑ_{200}
10	wind direction(200m)	θ_{200}
11	wind speed(500hPa)	ϑ_{500hPa}
12	wind direction(500hPa)	θ_{500hPa}
13	temperature(0m)	T_0
14	temperature(2m)	T_2
15	temperature(20m)	T_{20}
16	temperature(50m)	T_{50}
17	specific humidity (2m)	q
18	pressure(0m)	P
19	long wave down(0m)	Iw_d
20	short wave down(0m)	Sw_d
21	precipitation (0m)	p
22	wind speed(80m)	ϑ_{80}
23	wind direction(80m)	θ_{80}
24	power(80m)	P_w
25	ϑ_{10}^2	ϑ_{10}^2
26	ϑ_{10}^3	ϑ_{10}^3
27	$g(T_{20}/T_{50})$	$g_{T_{20/50}}$
28	$g(T_2/T_{50})$	$g_{T_{2/50}}$
29	$g(T_0/T_{50})$	$g_{T_{0/50}}$
30	$g(\vartheta_{500}/\vartheta_{50})$	$g_{\vartheta_{500/50}}$
31	$g(\vartheta_{500}/\vartheta_{20})$	$g_{\vartheta_{500/20}}$
32	$g(\vartheta_{500}/\vartheta_{10})$	$g_{\vartheta_{500/10}}$
33	$g(\vartheta_{50}/\vartheta_{20})$	$g_{\vartheta_{50/20}}$
34	$g(\vartheta_{50}/\vartheta_{10})$	$g_{\vartheta_{50/10}}$
35	$g(\vartheta_{20}/\vartheta_{10})$	$g_{\vartheta_{20/10}}$
36	u_{10}	u_{10}
37	v_{10}	v_{10}
38	u_{20}	u_{20}
39	v_{20}	v_{20}
40	u_{50}	u_{50}
41	v_{50}	v_{50}
42	u_{100}	u_{100}
43	v_{100}	v_{100}
44	u_{200}	u_{200}
45	v_{200}	v_{200}
46	u_{500hPa}	u_{500hPa}
47	v_{500hPa}	v_{500hPa}
48	u_{80}	u_{80}
49	v_{80}	v_{80}

Table 3.5: Results of the hourly and daily wind speed estimation by the ELM and MLR with all features considered (98).

	RMSE [m/s]	MAE [m/s]	r^2
hourly			
ELM	1.75	1.33	0.69
MLR	2.11	1.66	0.68
daily			
ELM	0.49	0.36	0.90
MLR	1.31	1.19	0.89

Table 3.6: Comparative best results of the hourly wind speed prediction by the ELM and MLR, with different fitness functions in the CRO and CRO-SL algorithms.

	CRO			CRO-SL		
FITNESS	RMSE [m/s]	MAE [m/s]	r^2	RMSE [m/s]	MAE [m/s]	r^2
ELM						
$f_1(\mathbf{x})$	1.72	1.32	0.70	1.67	1.27	**0.72**
$f_2(\mathbf{x})$	1.73	1.32	0.69	1.67	1.28	0.71
$f_3(\mathbf{x})$	1.72	1.32	0.70	1.66	1.27	0.71
MLR						
$f_1(\mathbf{x})$	1.80	1.37	0.69	1.73	1.33	0.70
$f_2(\mathbf{x})$	1.78	1.37	0.69	1.73	1.33	0.70
$f_3(\mathbf{x})$	1.77	1.35	0.69	1.72	1.32	0.70

Table 3.7: Best set of features selected by the CRO-SL ($f_1(\mathbf{x})$, 25 features).

Feature #	Variable
1	ϑ_{10}
2	θ_{10}
5	ϑ_{50}
7	ϑ_{100}
12	θ_{500hPa}
13	T_0
14	T_2
28	$g_{T_{2/50}}$
29	$g_{T_{0/50}}$
31	$g_{v_{500/20}}$
39	v_{10}
40	u_{50}
41	v_{50}
42	u_{100}
43	v_{100}
45	v_{200}
48	u_{80}
49	v_{80}
54	$ma(\vartheta_{50})$
56	$ma(\vartheta_{100})$
62	$ma(T_0)$
86	$ma(v_{10})$
88	$ma(v_{20})$
90	$ma(v_{50})$
98	$ma(v_{80})$

The effect of including several objective functions with different weights for hourly and daily prediction (see Equations (3.2) to (3.4)), can be evaluated by analyzing Table 3.8. Note that in daily prediction the inclusion of a feature selection mechanism also improves the performance of the prediction system (see also Table 3.5 to compare the results without feature selection). In this case the CRO-SL-ELM system is able to obtain a value for r^2 of 0.93, improving the 0.90 of the ELM without feature selection procedure. In spite of this improvement in prediction performance, note that f_1 is again the objective function which produces the best results in terms of daily wind speed prediction, which indicates that objective functions f_2 and f_3 (which included specific daily prediction terms), are not able to improve the prediction.

Another comparison is carried out by considering the best solutions found by the CRO-ELM and CRO-SL-ELM algorithms, and use them in alternative regression mechanisms. In this case, we consider the SVR [Smola2004] and the MLP neural network, trained with the Levenberg-Marquardt algorithm [Hagan1994], as final regression mechanisms. We have chosen these regressors since they have also provided good results in alternative problems. Table 3.9 shows the performance of both approaches using the solutions from the CRO-ELM and CRO-SL-ELM. As can be seen, the SVR is not able to improve the performance of the ELM in the final prediction. Moreover, it seems that the features selected by the CRO-ELM degrade the

Table 3.8: Comparative best results of the daily wind speed estimation by the ELM and MLR, with different fitness in the CRO-ELM and CRO-SL-ELM algorithm.

	CRO			CRO-SL		
FITNESS	RMSE [m/s]	MAE [m/s]	r^2	RMSE [m/s]	MAE [m/s]	r^2
ELM						
$f_1(\mathbf{x})$	0.48	0.36	0.92	0.47	0.37	**0.93**
$f_2(\mathbf{x})$	0.50	0.38	0.91	0.45	0.32	0.92
$f_3(\mathbf{x})$	0.51	0.37	0.90	0.47	0.36	0.92
MLR						
$f_1(\mathbf{x})$	0.67	0.47	0.89	0.56	0.38	0.89
$f_2(\mathbf{x})$	0.63	0.46	0.89	0.57	0.40	0.90
$f_3(\mathbf{x})$	0.65	0.45	0.88	0.56	0.40	0.90

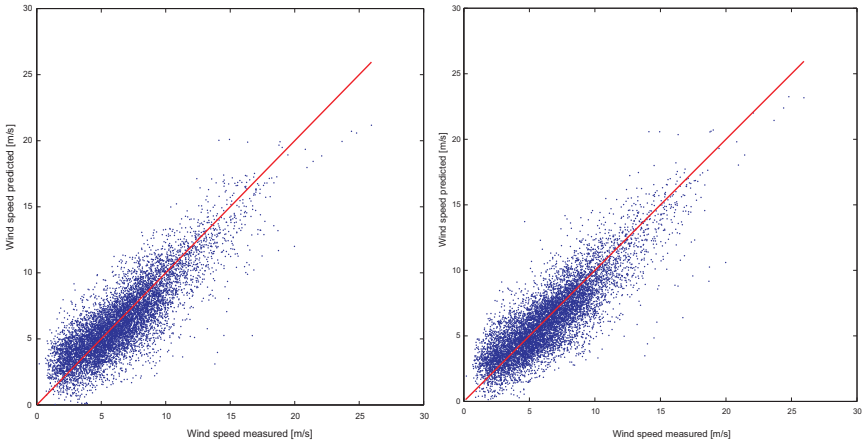
SVR performance significantly. The SVR using as inputs the solution by the CRO-SL-ELM performs better, but it is still not able to improve the prediction of the ELM as final regressor. This behaviour of the SVR can be determined by the number of features, since the CRO-ELM obtains a solution with 52, whereas the best solution by the CRO-SL-ELM has 25. The MLP has a better behaviour, producing better results than the SVR. The application of the feature selection mechanism seems to be positive in this case, and the MLP results improve when the CRO or the CRO-SL algorithms for FSP are previously applied. In this case, the system with the MLP as final regressor obtains better results than using the MLR, but it is worse than the CRO-SL-ELM previously analyzed.

Table 3.9: Comparative best results of the hourly wind speed estimation by a SVR and MLP regressors, with all features, CRO-ELM and CRO-SL-ELM, considering fitness function f_1.

	RMSE [m/s]	MAE [m/s]	r^2
SVR (All features)	2.82	2.13	0.20
SVR (FSP CRO)	2.63	1.97	0.40
SVR (FSP CRO-SL)	1.69	1.28	0.71
MLP (All features)	1.83	1.41	0.67
MLP (FSP CRO)	1.74	1.33	0.69
MLP (FSP CRO-SL)	1.71	1.30	0.70

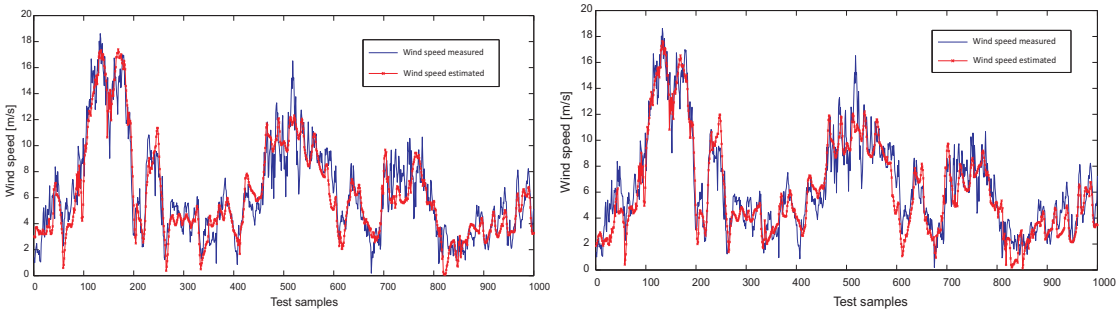
Figure 3.7 shows the scatter plot with and without feature selection (ELM and CRO-SL-ELM). This figure can be complemented with Figure 3.8, which shows the temporal performance of the predictions with and without feature selection mechanism. In general both predictions are

quite accurate, which shows the good performance of the ELM as regressor/predictor. Note that it is possible to detect differences due to the feature selection process, such as better wind speed peaks and valleys reconstruction, which produce the improvement in the prediction obtained with the CRO-SL-ELM algorithm.



(a) (b)

Figure 3.7: Scatter plots of the wind speed hourly estimation by the ELM method for the test data set: (a) without feature selection; (b) with CRO-SL for the selection.



(a) (b)

Figure 3.8: Temporal evolution of the wind speed hourly estimation by the ELM method for the test data set: (a) without feature selection; (b) with CRO-SL for the selection.

Chapter 4

Feature selection in a daily global solar radiation prediction system

4.1 Introduction

Solar energy is a clean and sustainable renewable source, with a high potential for significant growth in future years. Solar energy development is specially important in the Middle-East, southern Europe and the USA, places where the solar resource can be exploited all year around [Kalogirou2014]. An important problem faced by this resource is its integration in the grid system, because the energy produced by solar facilities is intrinsically stochastic due to the presence of clouds, atmospheric particles, dust, etc. In order to predict the solar production at photovoltaic facilities, an accurate GSR prediction at the solar plant is needed (in solar thermal systems, the direct solar radiation prediction is needed instead). In any case, the solar radiation reaching a point of the Earth surface completely depends on different atmospheric variables [Inman2013, Khatib2012, Voyant2011].

A vast amount of different AI-related techniques have been applied for the prediction of GSR [Voyant2017]. Most of them use different Soft-Computing techniques, with inputs based on meteorological and geographical parameters such as sunshine duration, air temperature, relative humidity, wind speed, wind direction, cloud cover, precipitation, etc [Rehman2008, Bilgili2011]. According to [Belu2013] the design, control and operation of solar energy systems requires long-term series of meteorological data such as solar radiation, temperature, or wind data. ANNs are one of the most applied methods in solar radiation prediction problems. In [Yadav2014] an exhaustive review on solar radiation prediction using ANNs is presented, describing forty two different researches, each one using as input variables several of the above-mentioned parameters. In [Mellit2008], the authors present different Soft-Computing techniques (ANN, Fuzzy logic, GAs, Expert systems, etc.) applied to different photovoltaic applications: sizing of PV systems, modeling, simulation and control of PV systems or prediction of PV production using atmospheric or meteorological data. In some of the works cited, meteorological and geograph-

ical parameters are also used to increase the accuracy of the systems implemented, such as in [Sozen2004], where ANNs are used with geographical parameters to estimate solar energy potential in Turkey. In [Behrang2010], different combinations of input variables are considered and tested with MLP and RBF neural networks. Moreover, results are then compared to conventional GSR prediction models, concluding that the ANN approaches perform better. In a similar research approach, [Dorvlo2002] carries out a comparison between MLP and RBF neural networks in a problem of solar radiation estimation. Experiments in eight stations in Oman show the good results obtained with the neural algorithms. The work in [Bou2017] presents the performance of ANN in solar energy prediction in Kuwait. Two different training approaches, gradient descent and Levenberg-Marquart algorithm are tested in five different Kuwaiti locations. Recently, the use of ELMs as a fast training method for ANNs has been applied due to the good GSR prediction results obtained. Sahin et al. [Sahin2013] apply ELMs using satellite measures, concluding that the ELM model performs better than ANN with back-propagation in terms of GSR estimation and computational time. The ELM's performance in alternative radiation prediction systems is also described in [Wu2016, Dong2014, Salcedo2013].

Alternative strong regression algorithms have been applied in solar radiation prediction problems. Support Vector Regression (SVR) is one of these approaches, which has been successfully exploited in solar radiation prediction. In [Mohammadi2015] the SVR algorithm has been mixed with a wavelet transform, in order to improve the performance of the former. Results in an Iranian coastal city have shown the performance of this hybrid proposal. The work [Olatomiwa2015] presents a firefly meta-heuristic with a SVR for GSR prediction in different locations of Nigeria. In [Antonanzas2015] solar irradiation mapping is tackled with SVR with exogenous data. Variable selection using a genetic algorithm is also considered in order to improve the performance of the SVR approach. In [Monteiro2017] a comparison between the performance of SVR and ANNs is carried out, in a problem of photovoltaic power generation. The work in [Belaid2016] tackles a problem of GSR prediction with SVRs considering different prediction time horizons. Finally, in [Chen2011, Qiao2013, Zengand2013], a least-square SVR has been implemented using meteorological and atmospheric data as predictive variables. The results obtained with the least-square SVR are compared with those of an auto-regressive neural network and a RBF neural network.

Bayesian methodology has also been applied in solar energy prediction problems. In [Lopez2005] the description of a Bayesian methodology to determine the most relevant meteorological input parameters for an ANN in order to determine direct solar irradiance is introduced. This study concludes that the clearness index and the relative air mass are the most significant ones. [Yacef2012] also discusses the performance of Bayesian networks in global solar radiation prediction. A comparison of GSR prediction performance among Bayesian networks, MLPs and empirical models is carried out. In [Wu2014], several clusters are formed and a prediction model is trained for each cluster to represent a different pattern in the stochastic component of the solar radiation, obtaining better results than ARMA or Time-Delay NNs. In [Salcedo2014c], an approach for daily global solar irradiation prediction based on temporal Gaussian processes is discussed, improving the performance of a number of alternative regressors such as NNs, SVR

or regression trees.

Other approaches use all-sky or satellite images in order to obtain solar radiation prediction, such as [Fu2013], where the prediction of the solar radiation is based on different features extracted from all-sky images, as previously discussed in Section 2.1.2. On the other hand, [Senkal2009] and [Deo2017] tackle the problem of solar radiation prediction from satellite images, in different locations in Turkey and Australia, respectively.

Hybrid approaches, i.e. algorithms which mix some kind of regression techniques with predictors from different sources have also been used in solar energy prediction problems. In [Bhardwaj2013] solar irradiation in India is analyzed using a hybrid approach that combines Hidden Markov Models and generalized fuzzy models. According to Diagne et al. [Diagne2013], forecasting of GHI can be categorized according to the input variables used, that also determine the forecast horizon where they perform best. For instance, for time horizons from 4 to 6 hours, the use of NWP models typically outperforms satellite-based predictions. Moreover, the use of the WRF meso-scale model (a regional NWP model) hybridized with a Kalman filter reduces de GHI hour-ahead forecast relative root mean square error (rRMSE) from 35.20% to 22.33% [Diagne2014]. In [Wu2011] a hybrid approach formed by Time Delay NNs and ARMA models is shown for short-term (hourly) solar radiation prediction in Singapore. A similar approach which also mixes ANN and ARMA models is proposed in [Voyant2013]. In this case, the performance of the methodology is evaluated in different location of the Southern French coast and Corsica. In [Hocaoglu2008] ANNs are mixed with 2-D linear filters to obtain a hybrid approach for hourly solar radiation prediction. In [Lima2016] ANNs are hybridized with numerical weather prediction models for solving a problem of surface solar irradiance prediction in Brazil. [Sharma2016] built a mixed approach formed by an ANN with wavelet transform for solving a problem of solar irradiance forecasting in 25 different locations of Singapore.

Finally, in [Aybar2016] a GGA is mixed with an ELM for solar radiation prediction at different time horizons. In that work, the GGA determines which WRF output variables best refine the GSR forecast (performed using the above-mentioned ELM approach). In [Salcedo2014b] a hybrid ELM-CRO is used for solar radiation prediction in Southern Spain, in which the CRO modifies the ELM weights in order to improve its performance.

In this chapter we use a hybrid CRO-SP with an ELM network approach to predict the GSR at a given location. For this purpose, we use a vast set of meteorological and atmospheric variables provided by the WRF, at different points close to the target location under study. Then, the CRO-SP algorithm is used to determine the best subset of WRF variables that lead to a best forecast. This problem is known in the literature as *statistically downscaling* the GSR prediction of a meso-scale model to a given point [Schmidli2007]. Therefore, the ultimate goal of this approach is to evaluate what features (predictive variables) from the numerical model are useful for this downscaling process. Note that in this chapter, each species in the CRO-SP algorithm represents the use of a different number of WRF variables, i.e. it encodes a different number of predictive variables considered in the prediction. The complete solar prediction system (CRO-SP-ELM) will be tested with real data from a radiometric station in Toledo, Spain.

4.2 Problem formulation

Let $\mathcal{I}_t$ be the global solar radiation in point $\mathcal{P}$ (a given location of the Earth's surface) at time t and let $\hat{\mathcal{I}}_t$ be the prediction of the global solar radiation in $\mathcal{P}$ at the same time instant t. Let $\mathcal{M}$ be a numerical meso-scale model and let $\mathcal{V}$ be the output of the model, at a time t at m different points of a grid. $\mathcal{V}$ consists of the prediction at time t of n different atmospheric variables, φ_{mn} ($m \in \{1, \ldots, M\}$ and $n \in \{1, \ldots, N\}$). Note that some or all of these variables may be registered at ground level ($l = 0$) or at different pressure levels ($l \in \{0 \ldots \mathcal{L}\}$). The output of $\mathcal{M}$ can be expressed as $\mathcal{V} = (\varphi_{11}, \ldots, \varphi_{1N}, \varphi_{21}, \ldots, \varphi_{2N}, \ldots, \varphi_{M1}, \ldots, \varphi_{MN})$, as shown in Figure 4.1.

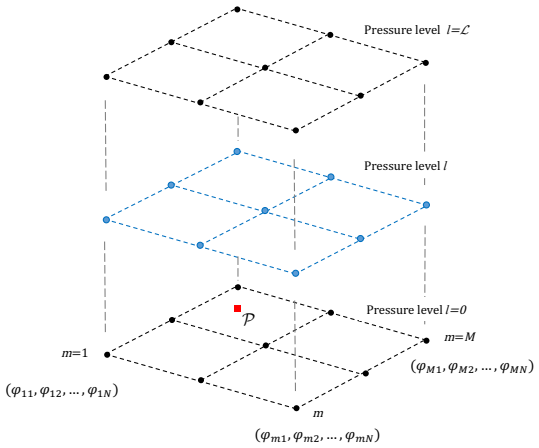


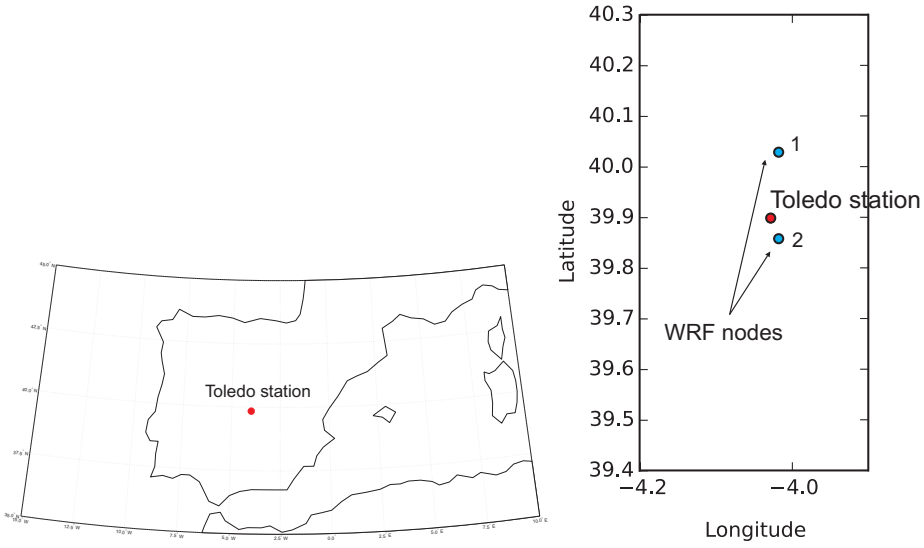
Figure 4.1: GSR prediction scheme used.

The problem we tackle consists in obtaining and accurate GSR estimation $\hat{\mathcal{I}}_t$ (in terms of the RMSE with respect to the real measurement $\mathcal{I}_t$) at point $\mathcal{P}$ from the WRF model outputs $\mathcal{V}$.

4.3 Objective variable data and predictive variables considered

In this work, $\mathcal{P}$ pinpoints at the radiometric station of Toledo, Spain (39° 53'N, 4° 02'W). Figure 4.2 (a) shows the measuring station's location within the Iberian Peninsula. The predictive variables considered $\mathcal{V}$ are the outputs of the WRF model at the two grid points ($M = 2$) closest to the station (in terms of the minimum Euclidean distance) and located at (39° 51'N,

4° 01'W) and (40° 02'N, 4° 01'W), respectively, see Figure 4.2 (b).



(a) (b)

Figure 4.2: Location of: (a) Toledo's measuring station in Spain and (b) the $M = 2$ WRF grid points considered for the downscaling

4.3.1 Objective variable data

The objective variable data to train and test the algorithms correspond to one year (from May 1st, 2013 to April 30th, 2014) of hourly global solar radiation data collected at Toledo's measuring station. This station is located at 39° 53'N latitude, 4° 02'W longitude and 515 MASL. Two constraints have been considered: 1) night hours present zero irradiance, and 2) a unique set of hours of interest, regardless of the season, is needed. Therefore, hourly data from 5 a.m. to 8 p.m. throughout the year are used in this analysis.

4.3.2 Predictive variables provided by the WRF

The WRF model outputs considered in the study are the following:

- OLR: Top of atmosphere outgoing long-wave radiation (W/m^2).

- GLW: Downward long-wave flux at ground level (W/m^2).

- SWDOWN: Downward short-wave flux at ground level (W/m^2).

- u: Zonal wind component at different pressure levels (m/s).

- v: Meridional wind component at different pressure levels (m/s).
- w: Vertical wind component at different pressure levels (m/s).
- PSFC: Atmospheric pressure at ground level (hPa).
- QVAPOR: Water vapor mixing ratio (in kg/kg). This variable is defined as the ratio of the mass of water vapor to the mass of dry air.
- TSK: Surface skin temperature (K).
- TH2: Potential temperature at 2 meters above the ground (K).
- T': Perturbation potential temperature (in K) at different pressure levels. The relationship between the perturbation potential temperature, T', and the potential temperature, θ, is $\theta = T' + 300$.
- CLDFRA: Total cloudiness (fraction of clouds in each cell) at different pressure levels. Cloud fraction ranges from 0 (no clouds) to 1 (clouds in a spatial grid cell).

Table 4.1 shows the 58 variables analyzed for each of the two grid point considered, indicating (when needed) the different pressure levels where they were obtained. A total of 116 variables are then considered in this study.

4.4 Methodology

As previously indicated, the CRO-SP algorithm is used to determine which set of WRF outputs obtains the best GSR prediction with an ELM. Details on these algorithms can be seen in Sections 1.2.3 and 1.2.1, respectively. Recall that each species in the CRO-SP algorithm represents the use of a different number of WRF variables, and we consider a binary encoding for this problem, so each species considers binary vectors with a different number of 1s on them. The health or fitness function considered for each coral (individual) is obtained computing the RMSE of the global solar radiation prediction. RMSE is used in this work instead of other validation metrics, because large forecast errors and outliers are weighted more strongly than smaller errors, as the latter are more tolerable in GSR prediction [Beyer2011, Kleissl2013].

To identify the best set and number of predictive variables, several experiments ($\mathcal{E}_i$, $i \in [1, \ldots, 3]$) have been run in a 10-fold cross validation scheme. Each CRO-SP experiment $\mathcal{E}_i$ consists of five sub-experiments, each one of them analyzing a specific species $\mathcal{S}_i$ ($i \in [1, \ldots, 5]$), i.e., all corals belonging to one species have the same number of features. The co-evolution of these species leads quickly to a coral-reef colonized by the most suited corals. Once convergence

Table 4.1: Outputs of the WRF model used in the experiments as predictive variables (58 variables per point of the WRF model).

variable	pressure levels (hPa)
OLR	-
GLW	ground
SWDOWN	ground
u	ground, 850, 700, 500, 400, 300, 200, 100, 50
v	ground, 850, 700, 500, 400, 300, 200, 100, 50
w	ground, 850, 700, 500, 400, 300, 200, 100, 50
PSFC	ground
QVAPOR	ground, 850, 700, 500, 400, 300, 200, 100, 50
TSK	ground
TH2	ground
T'	ground, 850, 700, 500, 400, 300, 200, 100, 50
CLDFRA	ground, 850, 700, 500, 400, 300, 200

is reached, the best coral in the reef belongs to a specific species and its health function stands for the RMSE value obtained in test. Note that to calculate the global solar radiation at each iteration, the ELM has been run 3 times and the health function value assigned to the coral is the average result obtained. In all the experiments, the CRO-SP algorithm has been run with the parameters shown in Table 4.2.

4.4.1 Results

Table 4.3 presents the results obtained in all the experiments carried out. The first experiment run, $\mathcal{E}_1$, is meant to resolve the order of magnitude of the number of features to be considered (10, 20, 30, 40 or 50), and it can be observed that the best prediction is found using 10 variables (RMSE = 68.21 W/m^2). Experiments $\mathcal{E}_2$ and $\mathcal{E}_3$ are used to refine the number of predictive variables to consider, both of them converging to best results when the species encode 8 variables.

Figure 4.3 presents the scatter plots for each experiments' best coral, showing the algorithm's good performance in all cases.

Figure 4.4 presents the comparison in time between the measured and the predicted GSR for experiment $\mathcal{E}_3$, the best one, where it can be seen that the prediction follows well the targeted series. Figure 4.5 shows the evolution with the number of iterations of the best coral in this

Table 4.2: CRO-SP optimization parameters.

Phase	Parameter
Inicialization	Reef size = 50×40 (2,000 positions) $\mathcal{S}_i$, $i \in \{1, \ldots, 5\}$ (5 species) $\rho_0 = 0.75$ (1,500 corals) $\rho_0^{\mathcal{S}_i} = 0.15$ (300 corals per species)
External sexual reproduction	$F_b = 0.70$ Random selection of broadcast spawners. Each possible coral must be broadcast spawner at least once per iteration k. New larva formation using 2-point crossover.
Internal sexual reproduction	$1 - Fb = 0.20$ $P_i = 0.30$
Larvae setting	$\eta = 3$ Identical corals are not allowed in the reef.
Asexual reproduction	$F_a = 0.05$ $P_a = 0.005$
Depredation	$F_d = 0.15$ $P_d = 0.25$ (it decreases with the number of iterations. At k_{max}, $P_d = 0$)
Stop criteria	$k_{max} = 300$ iterations.

Table 4.3: Experiments run considering different species. Each *species* is represented by $\mathcal{S}_i$.

Experiment	Number of features per species					RMSE (W/m^2)		Best Species
	$\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_3$	$\mathcal{S}_4$	$\mathcal{S}_5$	Average	Best coral	
$\mathcal{E}_1$	10	20	30	40	50	69.50	68.21	10 features ($\mathcal{S}_1$)
$\mathcal{E}_2$	6	8	10	12	14	69.33	68.16	8 features ($\mathcal{S}_2$)
$\mathcal{E}_3$	7	8	9	10	11	69.16	68.03	8 features ($\mathcal{S}_2$)

experiment. It corresponds to a coral encoding with 8 WRF variables and presents a final RMSE of 68.03 W/m^2 and a coefficient of determination r^2 =0.95.

It is interesting to analyze the evolution of the different species in the reef with the number of iterations. Figure 4.6 shows this evolution for the best run of the best experiment, $\mathcal{E}_3$. In Figure 4.6 (a) the random initialization of the reef is presented, where the reader can see the positions occupied by each different species and the free positions available at the reef. As the number of iterations (k) increases (Figures 4.6 (b)-(e)), it can be observed that the worst-fitted species tend to die and are no longer present at the reef, as larvae from dominant species outperform

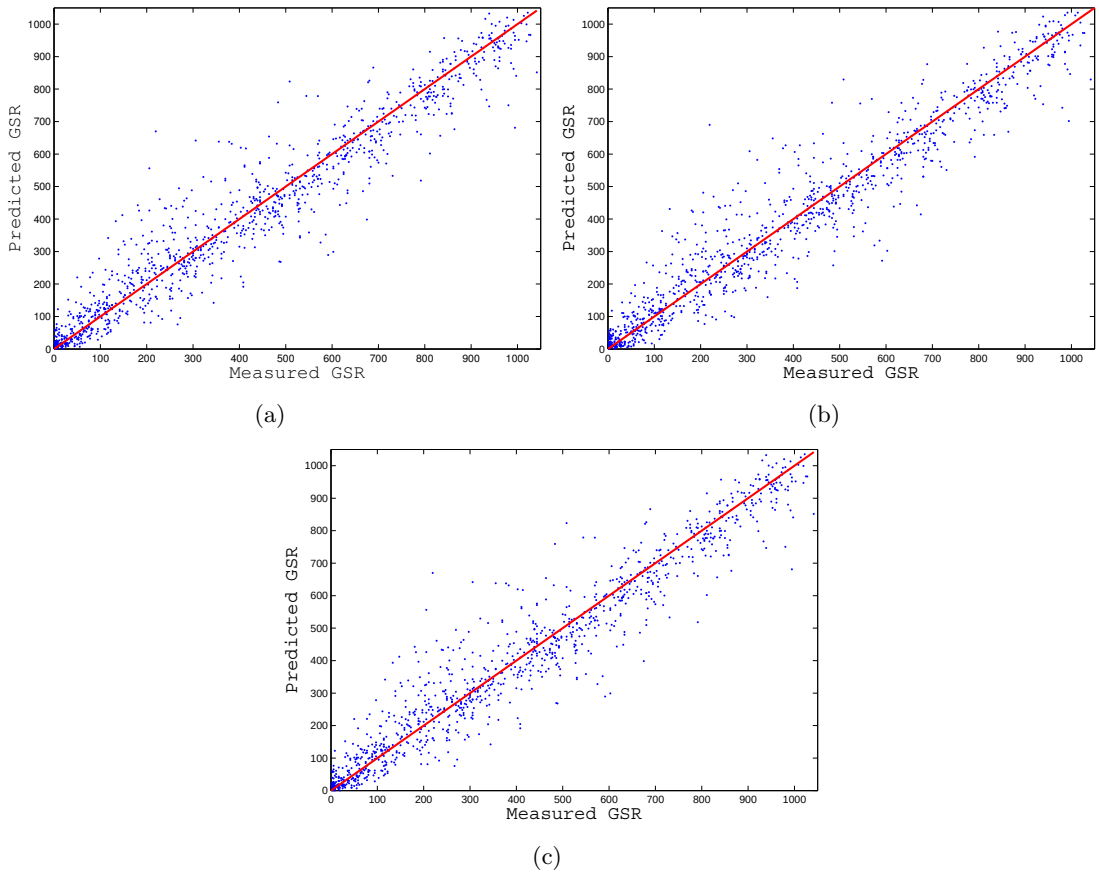
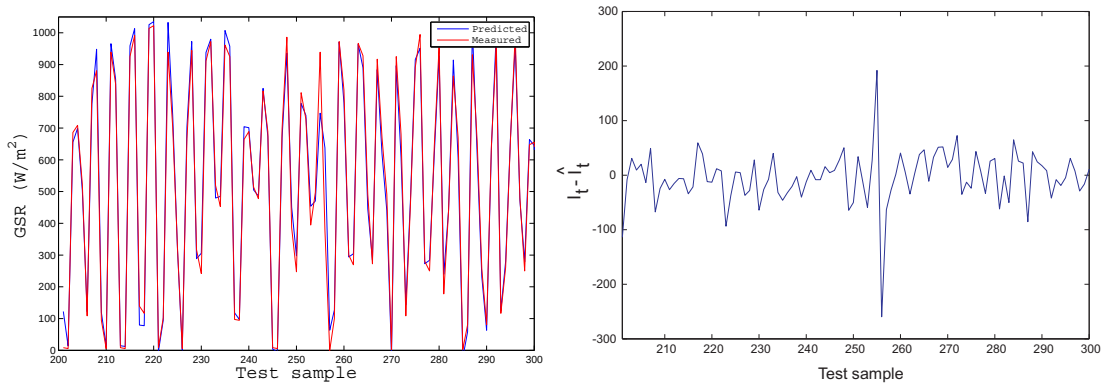


Figure 4.3: Scatter plot of the global solar radiation: (a) Experiment $\mathcal{E}_1$, (b) Experiment $\mathcal{E}_2$ and (c) Experiment $\mathcal{E}_3$.

them. Finally, when the stop criteria is reached, the reef is colonized by the best species which, in this particular experiment, is species $\mathcal{S}_2$ (corresponding to the use of 8 WRF variables for the prediction).

Figures 4.7 to 4.9 show, for all experiments analyzed, the evolution with the number of iterations of two important characteristics. First, the RMSE of each species' best coral, which is depicted in subfigures (a). It is clear that the RMSE decreases with the number of evolutions, but there is one exception: when a species is endangered (is being outperformed by the rest) its RMSE increases abruptly. Right after this occurs, the RMSE is interrupted, resulting in the disappearance of the worst-fitted species from the reef. Second, the number of corals present in each species is analyzed in Figures 4.7 (b), 4.8 (b) and 4.9 (b). It can be observed that, at some points, the number of corals in some species drops down. This is directly related to the occurrence of depredation phases. It is important to highlight that although in the depredation phase the species are decimated, the evolution keeps recovering the best-fitted.



(a) (b)

Figure 4.4: Experiment $\mathcal{E}_3$. (a) GSR over time. (b) Deviation in time of the predicted GSR from the measured GSR. Note that only a random time frame of 100 samples is presented for clarity purposes.

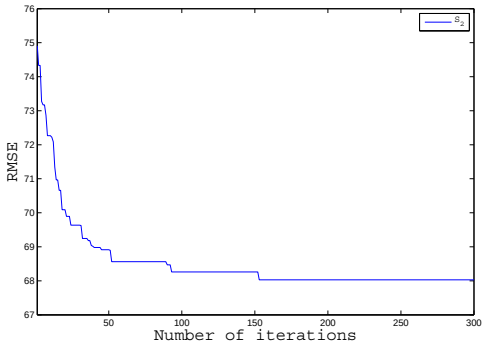
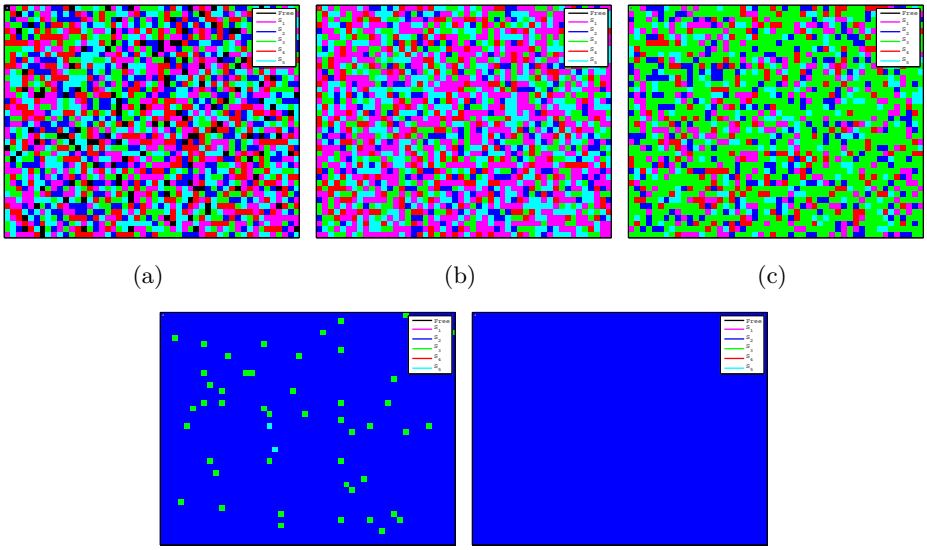


Figure 4.5: Experiment $\mathcal{E}_3$. Evolution with the number of iterations of the best coral's RMSE. Note that the best coral belongs to species $\mathcal{S}_2$.

Next, Table 4.4 shows the name of the best coral's WRF outputs selected for each experiment. It can be seen that there are six variables: OLR, w at 400 hPa, CLDFRA at 200 hPa, T at 850 hPa and T at 400 hPa corresponding to the first grid point, and v at 500 hPa corresponding to the second grid point, present in all experiments' results. Therefore, we can conclude that these variables set the rough prediction while the other WRF outputs perform the refinement. Thus, for the third experiment, the RMSE using these 6 variables (over the same test sets) is 74.05 W/m^2 and 72.59 W/m^2, average and best values respectively. Once the refinement takes place, these RMSE values drop down to 69.16 W/m^2 and 68.03 W/m^2 respectively (as stated in Table 4.3).

Finally, in Table 4.5 the results are compared to those obtained with other techniques. First,



(a) (b) (c)

(d) (e)

Figure 4.6: Experiment $\mathcal{E}_3$. Evolution of the species present in the reef after a certain number of iterations (k): (a) $k = 1$; (b) $k = 10$; (c) $k = 25$; (d) $k = 50$; (e) $k = 150$. Each color pixel stands for a different coral species and free cells in the reef (free (black), $\mathcal{S}_1$ (magenta), $\mathcal{S}_2$ (blue), $\mathcal{S}_3$ (green), $\mathcal{S}_4$ (red) and $\mathcal{S}_5$ (cyan)).

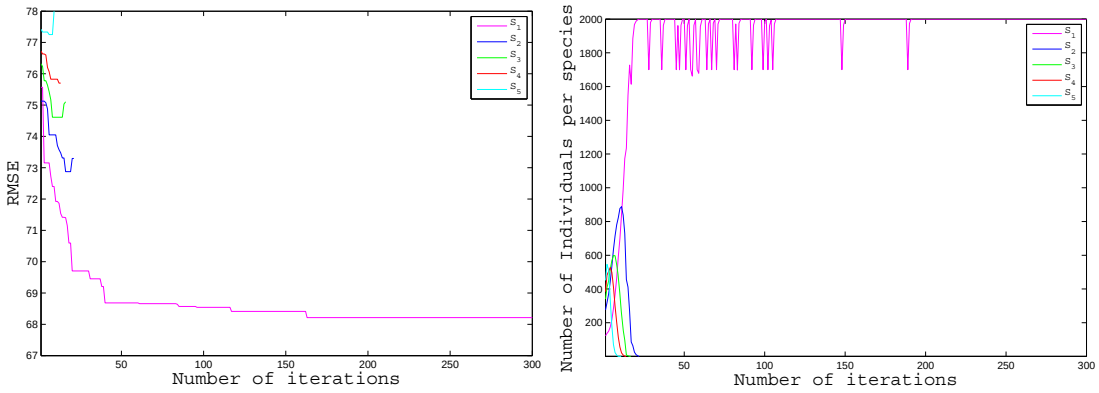


Figure 4.7: Experiment $\mathcal{E}_1$. Evolution with the number of iterations of: (a) The RMSE of each species' best coral, and (b) The number of corals per species.

the reader can see the GSR prediction using the 116 WRF variables (no feature selection) as inputs to the ELM. Then, feature selection is performed using three different techniques: a GA, a GGA (as described in [Aybar2016]) and the proposed CRO-SP approach, and the variables

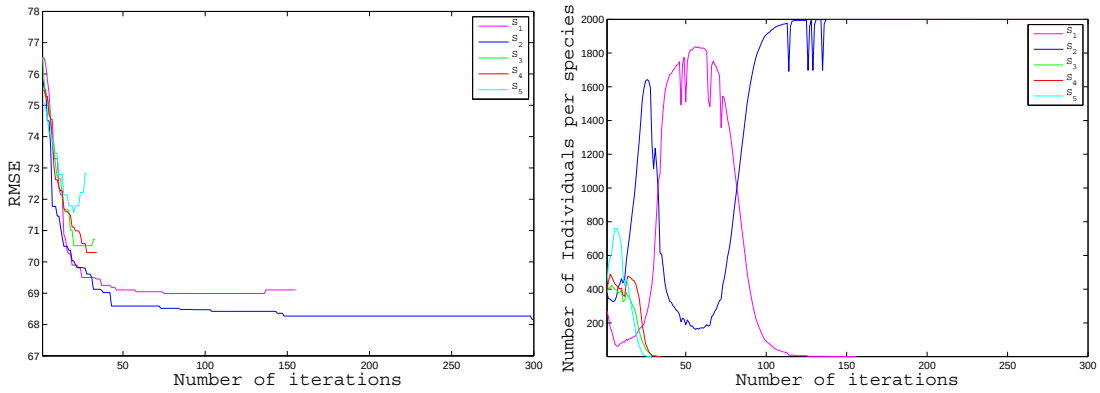


Figure 4.8: Experiment $\mathcal{E}_2$. Evolution with the number of iterations of: (a) The RMSE of each species' best coral, and (b) The number of corals per species.

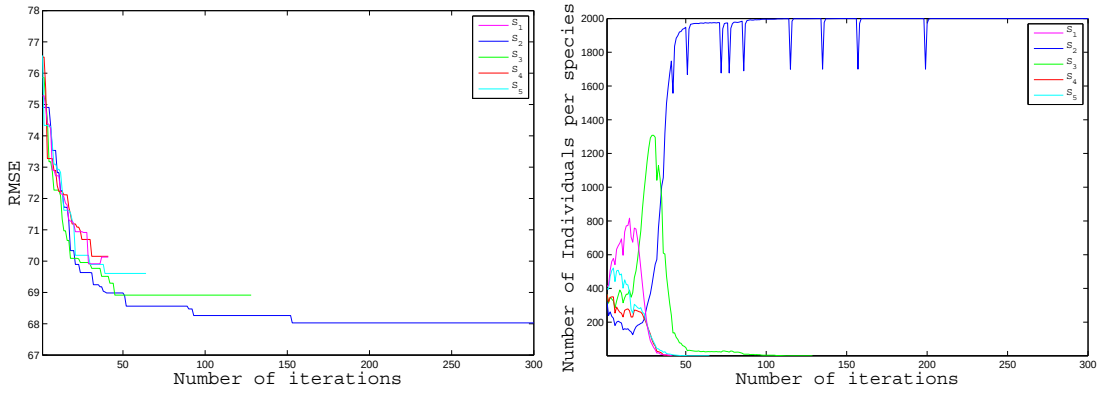


Figure 4.9: Experiment $\mathcal{E}_3$. Evolution with the number of iterations of: (a) The RMSE of each species' best coral, and (b) The number of corals per species.

chosen are used as the inputs to the ELM. It can be seen that the best results are obtained when the CRO-SP is used.

Table 4.4: Best predictive variables found for each experiment. Those variables present in all three experiments' results have been highlighted in bold face.

Experiment	Best WRF outputs selected	
	Grid point #1	Grid point #2
$\mathcal{E}_1$	**OLR**, ***w* (400 hPa)**, **CLDFRA (200 hPa)**, **T (850 hPa)**, **T (400 hPa)**, u (100 hPa)	***v* (500 hPa)**, PSFC, TH2, u (50 hPa)
$\mathcal{E}_2$	**OLR**, ***w* (400 hPa)**, **CLDFRA (200 hPa)**, **T (850 hPa)**, **T (400 hPa)**,	***v* (500 hPa)**, TH2 w (300 hPa)
$\mathcal{E}_3$	**OLR**, ***w* (400 hPa)**, **CLDFRA (200 hPa)**, **T (850 hPa)**, **T (400 hPa)**,	***v* (500 hPa)**, u (850 hPa) u (50 hPa)

Table 4.5: Comparison of the results obtained with other metaheuristic techniques.

Metaheuristic technique	RMSE (W/m^2)	
	Average	Best individual
ELM (No feature selection)	88.24	87.25
GA-ELM	73.98	72.20
GGA-ELM [Aybar2016]	74.73	73.66
CRO-SP-ELM	69.16	68.21

Chapter 5

Representative selection for robust temperature fields reconstruction

5.1 Introduction

The last years witnessed on-going growing interest in data-driven methods for climate sciences. Part of this interest is due to the large amount of new freely available data sources. However, a huge amount of data or information sources may have negative impact in the production of competitive results, when the methods for information processing are not fast or scalable enough. The improvement of computational approaches in the last years has been important, but in some cases it is very difficult to correctly process huge datasets with highly heterogeneous information. In these cases, it is important to use alternative information processing approaches which help reduce the computational load of a given problem, while keeping as much as the information of the dataset. There are different examples of this in climate sciences, such as problems of Feature Selection (previously described in this Thesis) for improving prediction systems, or Clustering approaches [Nasseri2011, Coggins2014, Jinming2015], with a large variety of specific applications.

Another important approach related to the reduction of computational complexity in data is the RS problem (see Section 1.4 for a detailed description of the problem and the associated AM). Recall that the RS is the problem of finding exemplar samples from a given data, points or items collection, in such a way that the selected exemplars accurately summarize the complete set of starting data [Wang2017]. In this chapter we focus on RS problems for field reconstruction of monthly average temperature, defined by time series obtained in a large number of measuring points or stations. The idea is to obtain a reduced number of representative measuring points or stations, in such a way that a reconstruction algorithm works as accurately as possible, in terms of a given error measure. In this case, we have chosen the AM [Lorentz1969] as reconstruction algorithm for the field (see Section 1.4 for details). The reconstruction error acts in this case as a good alternative measure for the information lost in the RS process, i.e. the smaller the

reconstruction error from the representative measuring points, the smaller the information lost in the RS process. Note that the optimization problem associated with the RS is hard, since the objective function is highly non-linear, and it has not a direct mathematical representation (it is a black-box from the AM output). In this work we explore the optimization capacities of the CRO-SL algorithm in the RS problem for temperature field reconstruction. Experiments in two different datasets, European Climate Assessment & Dataset [Klein2002] and ERA-Interim reanalysis [Dee2011] in Europe show the effectiveness of the proposed approach.

5.2 CRO-SL for RS in temperature fields reconstruction

The RS can be tackled as a pure integer optimization problem, where the objective function is to minimize the RMSE provided by the AM reconstruction algorithm. In this subsection we discuss the chosen problem encoding, and then we describe the different substrate layers defined in the CRO-SL approach.

5.2.1 Problem encoding in the CRO-SL

One of the key points in the definition of any optimization method is the encoding to face a given problem. In this case, the optimal location of representative nodes in temperature fields reconstruction requires the selection of a measuring points subset, out of an initial complete set. It is therefore a discrete problem, in which the encoding must indicate what points are selected. Different encodings can be applied for this problem, such as binary vectors $\mathbf{x} \in \mathbb{B}$, where $x_i = 1$ stands for the selection of the measuring point i, whereas $x_k = 0$ means that the point k has not been selected. This encoding has the advantage that it admits specific operators devoted to binary vectors. On the other hand, the individuals obtained are long (the individual length is in this case equal to $|S|$), and the number of 1s must be controlled in order to provide feasible solutions, i.e. the corals in the CRO-SL meta-heuristic must be corrected [Salcedo2009c]. Another possibility is to manage integer vectors as solutions, $\mathbf{x} \in \mathbb{N}$. In this case, the encoding is shorter than in the binary encoding, since the length of each individual is now equal to N, but it is necessary to control that there are not repeated measuring points in the individual for it to be feasible. Finally, we have chosen this latter encoding, because it provides a more compact version of the algorithm, and allows a better implementation of some of the searching procedures such as the HS. Figure 5.1 shows an example of the encoding used, and how it is translated into a solution in the CRO-SL algorithm.

5.2.2 Substrates considered in the CRO-SL

The considered substrates for solving the problem at hand are detailed below. Note that there are general purpose substrates, such as DE or HS-based, and other specific substrates with crossovers and mutations adapted to the problem at hand, and specifically to the chosen

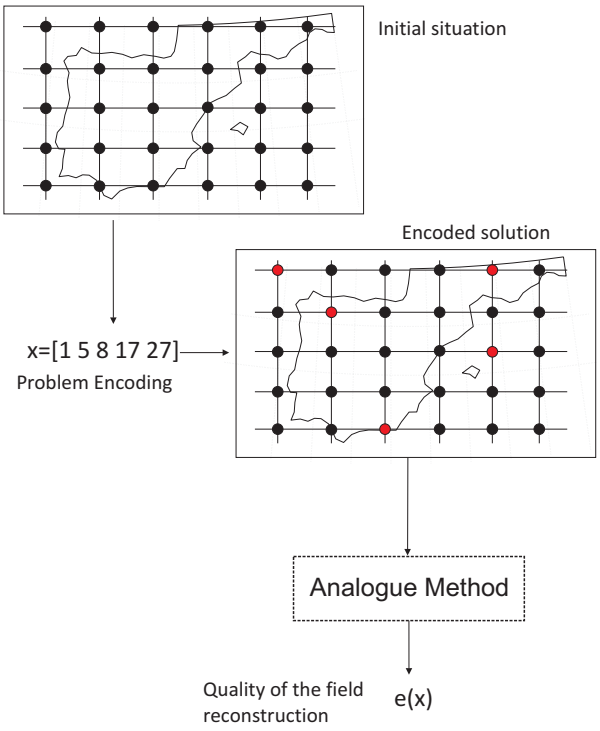


Figure 5.1: An example of the integer encoding of solutions chosen for the RS problem at hand in the CRO-SL.

encoding. A total of 5 substrates will be described and evaluated later in the experimental section.

- **Differential Evolution-based operator** (DE): This operator is based on the evolutionary algorithm with that name [Storn1997], a method with powerful global search capabilities. DE introduces a differential mechanism for exploring the search space. Hence, new larvae are generated by perturbing the population members using vector differences of individuals. Perturbations are introduced by applying the rule $x'_i = x^1_i + F(x^2_i - x^3_i)$ for each encoded parameter on a random basis, where x' corresponds to the output larva, x^t are the considered parents (chosen uniformly among the population), and F determines the evolution factor weighting the perturbation amplitude.

- **Harmony Search-based operator** (HS): Harmony Search [Geem2001] is a population based MH that mimics the improvisation of a music orchestra while composing a melody. This method integrates concepts such as harmony aesthetics or note pitch as an analogy for the optimization process, resulting in a good exploratory algorithm. HS controls how new larvae are generated in one of the following ways: i) with a probability HMCR$\in [0, 1]$ (Harmony Memory Considering Rate), the value of a component of the new larva is drawn

uniformly from the same values of the component in the other corals. ii) with a probability PAR$\in [0, 1]$ (Pitch Adjusting Rate), subtle adjustments are applied to the values of the current larva, replaced with any of its neighboring values (upper or lower, with equal probability).

- **Two points crossover** (2Px): 2Px [Eiben2003] is considered one of the standard recombination operators in evolutionary algorithms. In the standard version of the operator, two parents from the reef population are provided as input. A recombination operation from two larvae is carried out by randomly choosing two crossover points, interchanging then each part of the corals between those points.

- **Multi-points crossover** (MPx): Similar to the 2Px, but in this case the recombination between the parents is carried out considering a high number of crossover points (M), and a binary template which indicates whether each part of one parent is interchanged with the corresponding of the other parent.

- **Standard integer Mutation** (SM): This operator consists of a standard mutation in integer-based encodings. It consists of mutating an element of a coral with another valid value (different from the previous one). Note that the SM operator links a given coral (possible solution) to a neighborhood of solutions which can be reached by means of a single change is an element of the coral.

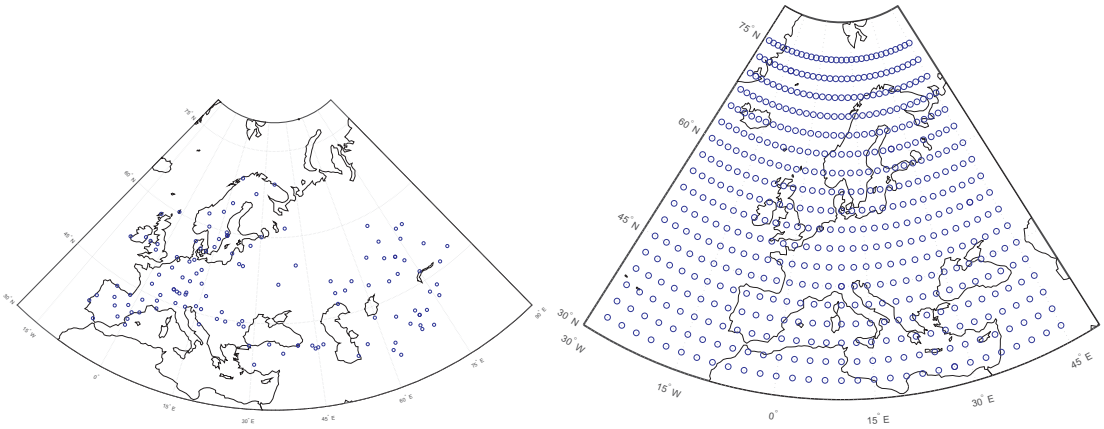
5.3 Experimental evaluation

This section describes the different experiments carried out to evaluate the CRO-SL performance in this RS problem. We have structured the section into several subsections: first, Section 5.3.1 describes the temperature datasets used in the experimental evaluation, which include gridded and un-gridded time series of temperature in Europe and central Asia. Then, we introduce a first subsection with fully resoluble situations, in which brute force algorithms can obtain the optimal solution. This section helps understand the difficulty of the problem, and introduces a constructive approach for comparison (the greedy algorithm for measuring points selections). Section 5.3.3 summarizes the performance evaluation of the CRO-SL in the problem. Discussions on the consistency of the algorithm, evaluation of the temperature field reconstruction obtained and a note on the computational performance of the CRO-SL are described in further subsections.

5.3.1 Temperature datasets used

The first dataset we consider to illustrate the performance of the CRO-SL algorithm in the temperature RS problem, consists of un-grided monthly average temperature data from the European Climate Assessment & Dataset (ECA) [Klein2002] and from several measuring

stations from the European HISTALP project [Chimani2013], which includes Central Europe and the Alps. The final dataset (ECA dataset hereafter) consists of 123 stations ($|S| = 123$), with monthly average temperature data, measured between January 1940 and December 2010. A total of $\hat{t} = 864$ months are then available, and we have divided them into training period ($t^T = 432$) months and test period ($t^V = 432$) months. This dataset has been previously described in a different study [Chidean2015], where the data treatment for eliminating the gaps in the series is described. Figure 5.2 (a) shows the location of the 123 measuring stations of the ECA case. The second dataset considered are gridded data from the ERA-Interim reanalysis of the ECMWF [Dee2011]. In this case, monthly average temperatures from January 1979 until July 2017 are taken into account (ERA dataset hereafter). A total of $\hat{t} = 462$ months are then available, and we have divided them into training ($t^T = 231$) and test ($t^V = 231$ months) periods. Figure 5.2 (b) shows the location of the reanalysis nodes ($|S| = 540$).



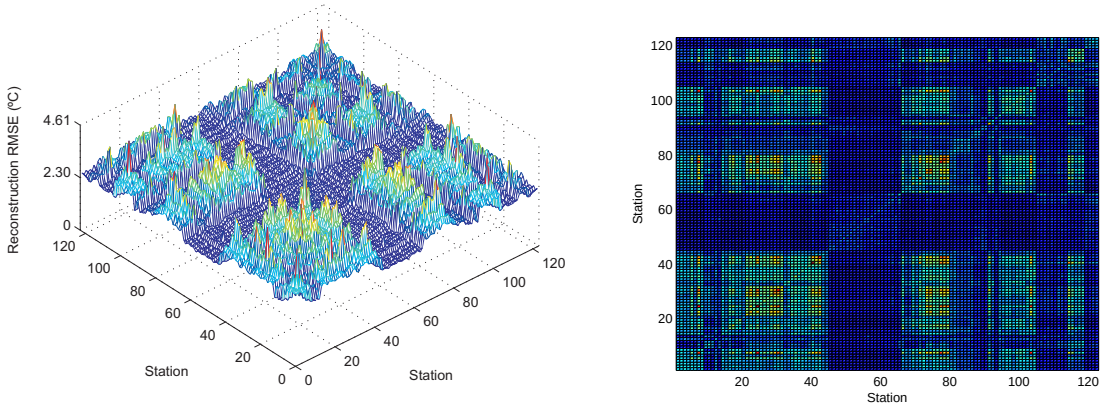
(a) (b)

Figure 5.2: Location of measuring stations (ECA) and reanalysis points (ERA); (a) Measuring stations of ECA and HISTALP dataset (un-grided measuring stations) for the first temperature RS problem considered; (b) Location of the ERA-Interim reanalysis nodes considered (gridded data), the second dataset in this study.

5.3.2 Results I: Case $N = 2$

First, we describe as special case, $N = 2$, a fully resoluble situation. Note that the optimal solution for this case can be obtained by calculating the RMSE in the field reconstruction for all possible pairs of stations (i, j) where $i, j \in \mathbf{s}$, and such that $i \neq j$ ($i = j$ is the $N = 1$ case, also described later in this subsection), i.e. a *brute force* computation algorithm. Figure 5.3 (a) shows the complete fitness landscape for the ECA dataset, which results in a discrete surface of 123×123 points. Figure 5.3 (b) shows a two-dimensional view depict from above. In these figures it is possible to see how different pairs of selected measuring stations lead to

different reconstruction error. In other words, there are pairs of stations which better represent the temperature field than others, in terms of getting a better reconstruction error from them by applying the AM. In Figure 5.3 (a) It is possible to locate large zones of error peaks and valleys, with a minimum value of the objective function $e(\mathbf{s}_N) = 2.05$ °C, with $\mathbf{s}_N = [19, 58]$, i.e., the best representative pair of stations in terms of reconstruction error are stations 19 and 58.

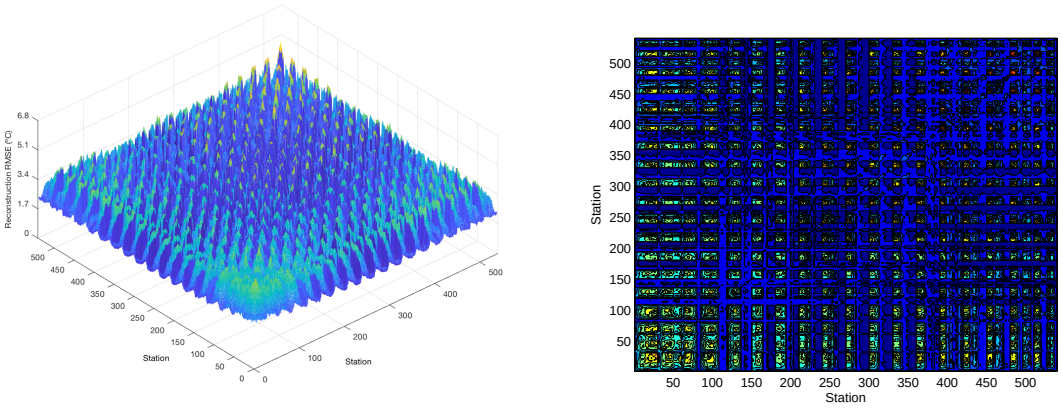


(a) (b)

Figure 5.3: RMSE landscape in the ECA dataset, for all combinations (pairs) of measuring stations (special case $N = 2$); (a) Landscape (3D); (b) Contour plot (2D).

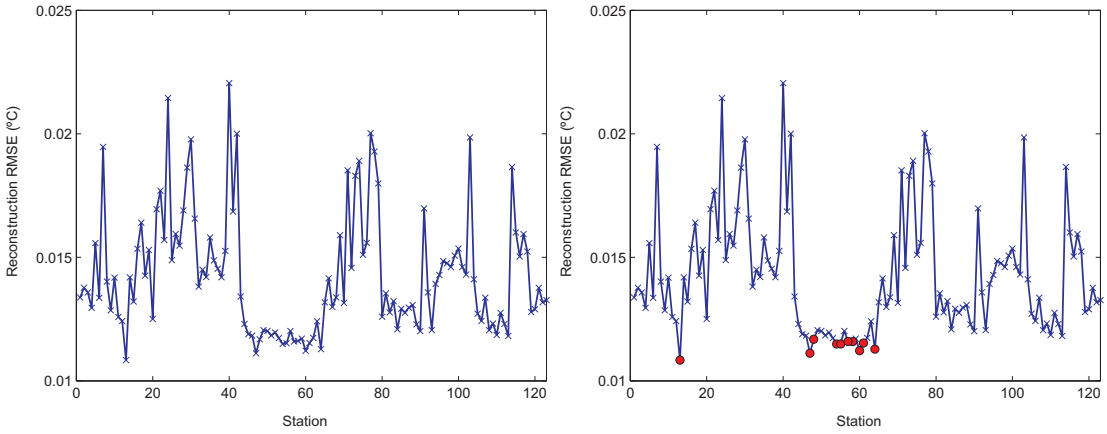
The error reconstruction landscape for the $N = 2$ case in the ERA dataset is shown in Figure 5.4 (a), with the two-dimensional graph from above shown in Figure 5.4 (b). In this case the graph shows a discrete surface of 540×540 measuring nodes, where it is also possible to detect peaks and valleys of RMSE, with the same conclusion as before: there are some pairs of measuring nodes which are better than others for representing the complete field of temperature. In this data set, the minimum value of the objective function is $e(\mathbf{s}_N) = 1.76$ °C, where $\mathbf{s}_N = [65, 264]$.

Let us finally discussing the case $N = 1$, for the ECA dataset. Figure 5.5 (a) shows the curve of reconstruction error using the AM when considering $N = 1$ station to set the best reconstruction time. Note that this curve is the main diagonal of the reconstruction error matrix depicted in Figure 5.3 (a). It is straightforward that there are stations which represent better the temperature field than others. In addition, it is possible to obtain a *greedy* (constructive) algorithm from this simple case, in the following way: sequentially add the measuring stations or points which provide the best solution for the simplest case $N = 1$. For example, Figure 5.5 (b) shows the construction of a solution for the case $N = 10$ by sequentially adding the best stations in the $N = 1$ reconstruction error graph. The final solution is $\mathbf{s}_N$=[13 ,47, 60, 64, 54, 55, 61, 57, 58, 48] with a fitness $e(\mathbf{s}_N) = 2.09$ °C.



(a) (b)

Figure 5.4: RMSE landscape in the ERA dataset, for all combinations (pairs) of measuring stations (special case $N = 2$); (a) Landscape (3D); (b) Contour plot (2D).



(a) (b)

Figure 5.5: Best representative station (case $N = 1$) for the ECA dataset and greedy approach construction; (a) Reconstruction error for ($N = 1$); (b) Greedy (construction based) solution obtained for the case $N = 10$.

5.3.3 Results II: general experimental performance

In this subsection we present the results obtained when applying the CRO-SL algorithm to select the best N representative measuring stations or points in the ECA and ERA datasets. Table 5.1 shows the parameters of the algorithm in all the experiments carried out in the paper. Four different cases are taken into account in both datasets: $N = 5, 10, 15$ and 20. Note that they cannot be solved by brute force algorithms, so meta-heuristics such as the proposed CRO-SL approach are a very good option to solve the problem. Table 5.2 shows a summary of the

Table 5.1: CRO-SL optimization parameters.

Phase	Parameter
Initialization	Reef size = 50×40 (2,000 positions)
	$\rho_0 = 0.9$
External sexual reproduction	$F_b = 0.80$
	$\mathcal{T} = 5$ substrates: HS, DE, 2Px, MPx, SM
Internal sexual reproduction	$1 - Fb = 0.20$
Larvae setting	$\kappa = 3$
Asexual reproduction	$F_a = 0.05$
Depredation	$F_d = 0.15$
	$P_d = 0.05$
Stop criterion	$k_{max} = 1000$ iterations.

results obtained by the CRO-SL algorithm, and a comparison with alternative approaches for this problem: a HS algorithm [Geem2001], a DE approach [Storn1997], a Hill Climbing heuristic [Michalewicz2000] and the greedy approach described in the previous section. This table shows the RMSE in °C, for the field reconstruction with the AM for the best solution found by the CRO-SL, following Equation (1.6). The average and standard deviation (out of 5 different runs of the algorithm) are also shown. It is possible to see how the reconstruction error is better when the number of representative stations or points grows. Note that we can compare the figures in Table 5.2 with the reconstruction with the maximum information possible ($N = |S| = 123$ for the ECA dataset and $N = |S| = 540$ for the ERA dataset). In the ECA dataset, the RMSE for the field reconstruction using the AM is 1.43 °C, whereas in the ERA dataset it is 1.31 °C. Thus, the results obtained with a reduced number of stations/nodes selected by the CRO-SL (representative stations/nodes) are close to the best possible result with all the available information. Regarding the comparison with alternative algorithms, first, note that the results obtained by the greedy algorithm are far away from those by the other methods tested, in terms of the RMSE. This approach constructs the final solution from the best set of N measuring stations/points considering only 1 station. The failing of this procedure means that the information of single stations is not enough to obtain a good set of N measuring points, so complete solutions of N stations must be looked for. Regarding the comparison with other heuristic and meta-heuristic algorithms, the CRO-SL is able to obtain the best results consistently, both in the ECA and ERA datasets. The second best algorithm is in this case the HS, which performs similar in the smallest cases ($N = 5$ and $N = 10$), but the differences in the largest instances $N = 15$ and $N = 20$ are higher. The DE and Hill Climbing approaches do not obtain results as competitive as the CRO-SL algorithm in any tested case.

The best solutions found by the CRO-SL in the ECA and ERA cases are displayed in Figures 5.6 and 5.7, respectively. It is possible to see how the CRO-SL obtains solutions which maximize the coverage of the region considered, as could be expected from a representative selection point

Table 5.2: Results in terms of the RMSE in the field reconstruction (in °C), obtained in ECA and ERA datasets, by the CRO-SL, HS, DE, Hill Climbing and greedy approaches.

	CRO-SL (ECA)			CRO-SL (ERA)		
	Best	Mean	Var	Best	Mean	Var
$N = 5$	1.67	1.67	0	1.49	1.51	0.0011
$N = 10$	1.53	1.53	0.0031	1.38	1.41	0.0067
$N = 15$	1.48	1.49	0.0038	1.34	1.37	0.0038
$N = 20$	1.45	1.45	0.0018	1.33	1.34	0.0044
	HS (ECA)			HS (ERA)		
$N = 5$	1.67	1.67	0	1.49	1.50	0.0016
$N = 10$	1.54	1.54	0	1.39	1.39	0.0004
$N = 15$	1.50	1.50	0.0021	1.37	1.37	0.0107
$N = 20$	1.48	1.48	0.0029	1.35	1.36	0.0052
	DE (ECA)			DE (ERA)		
$N = 5$	1.71	1.71	0.0048	1.52	1.52	0.0019
$N = 10$	1.58	1.58	0.0019	1.43	1.43	0.0009
$N = 15$	1.51	1.52	0.0030	1.39	1.39	0.0087
$N = 20$	1.48	1.49	0.0039	1.37	1.37	0.0059
	Hill Climbing (ECA)			Hill Climbing (ERA)		
$N = 5$	1.77	2.04	0.0350	1.58	1.90	0.0664
$N = 10$	1.61	1.76	0.0083	1.46	1.63	0.0116
$N = 15$	1.54	1.66	0.0042	1.41	1.54	0.0053
$N = 20$	1.53	1.60	0.0022	1.39	1.50	0.0030
	Greedy (ECA)			Greedy (ERA)		
$N = 5$	2.14	-	-	1.80	-	-
$N = 10$	2.10	-	-	1.71	-	-
$N = 15$	2.06	-	-	1.68	-	-
$N = 20$	2.02	-	-	1.67	-	-

of view. For a reduced number of stations or nodes (low values of N), the algorithm must choose those stations/nodes with the most possible information for the AM. In the ECA dataset the stations which are considered as the representatives cover central Europe, Scandinavia, Central Asia, Black Sea and Southern Asia. When $N = 10$, Europe is modelled with two representative stations, one for central Europe and another one for southern Europe located in Spain. The Scandinavian station is kept, and another close station is also selected in northern Russia. The Black Sea is now covered with two stations, one in the East and another one in the West, and the rest of Asia is covered by four stations, two of them in similar locations as in the $N = 5$

case. When N grows, the representative stations still cover the same areas, but in this case with increasing density, resulting in a more accurate temperature field reconstruction by applying the AM.

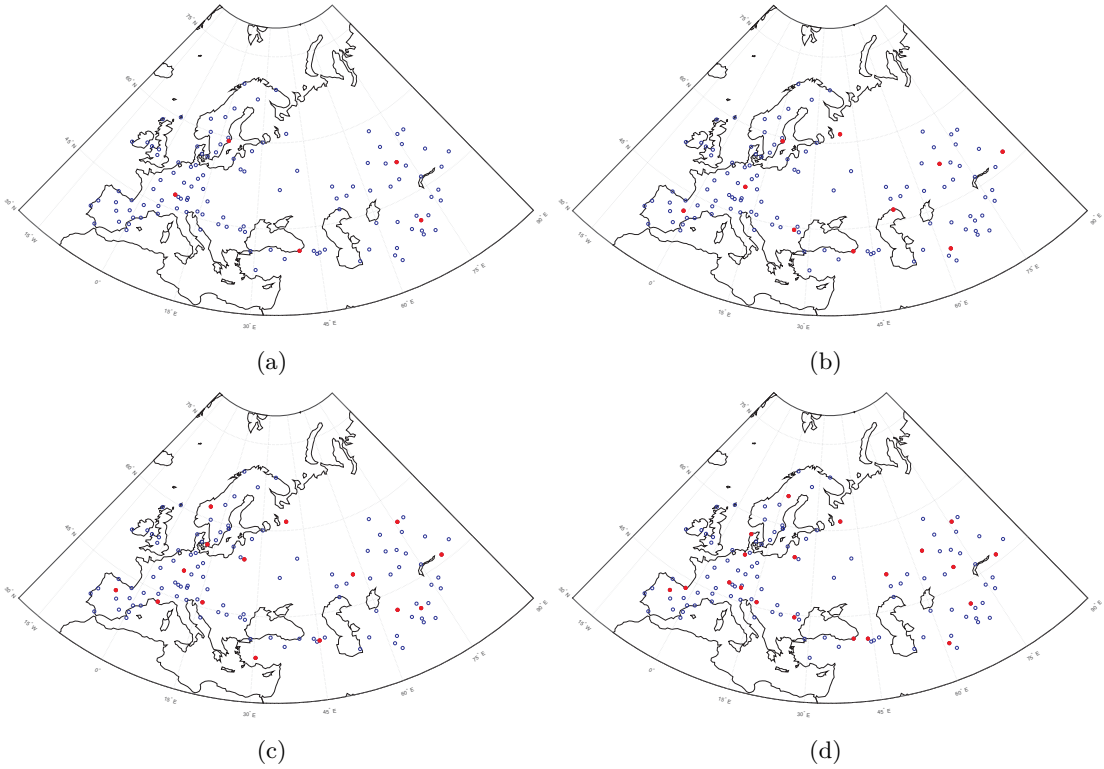


Figure 5.6: Best solution found by the CRO-SL (red points stand for the selected representative measuring stations), for the ECA dataset; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

The ERA dataset case is somehow different, since now the total area to be covered by the representative measuring points is smaller, there are more nodes to be selected from, and sea nodes are available (in the ECA dataset we only have inland measuring stations). When $N = 5$ is considered, one node is located at the north of Iceland, to cover polar regions, another node is located in Scandinavia, a node in northern Italy covers central and southern Europe and the two remaining nodes cover eastern Europe and the Mediterranean with a node in Egypt. When $N = 10$, two nodes cover the polar region, three nodes are devoted to Scandinavia, two nodes cover central and eastern Europe, one node at the south of Spain covers southern Europe and there are other two nodes, for the Mediterranean and Black Sea regions. The solutions with larger N follow this trend, by locating more representative nodes in the pole and Scandinavia, selecting some of them for central and eastern Europe, and locating the rest in the Mediterranean and southern Atlantic zones.

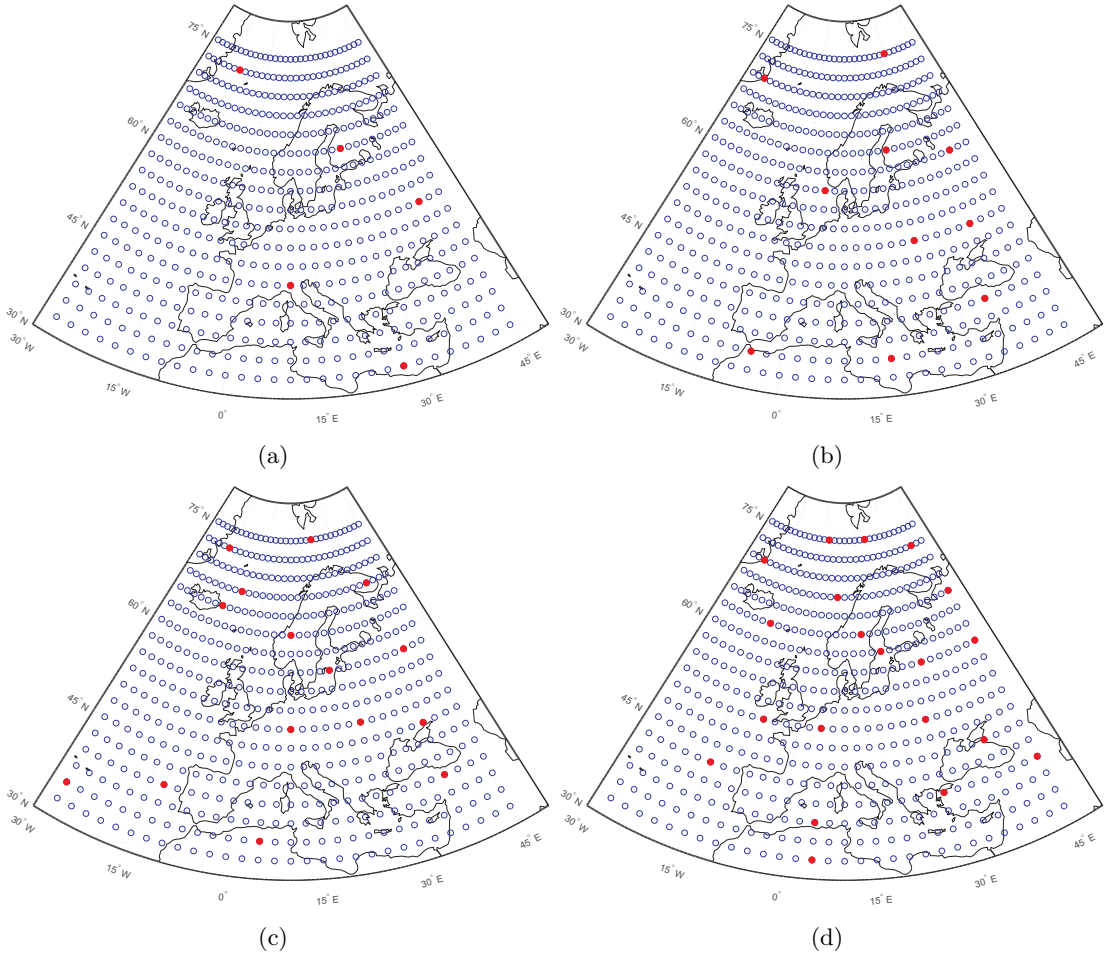


Figure 5.7: Best solution found by the CRO-SL (red points stand for the selected representative nodes), for the ERA dataset; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

5.3.4 Discussion I: consistency of the CRO-SL performance and the results obtained

In order to check the consistency of the results obtained by the CRO-SL, we have tackled the *hindcasting* problem, i.e. interchanged the Train period by the Test period in both datasets. This way, now the training is carried out with the last period in the dataset, and the test is over the earlier one. Table 5.3 shows the results obtained. In general, the results are slightly worse in terms of the reconstruction error than in the previous version of the problem. However, these minor differences are present for all values of N used in the CRO-SL and greedy approaches, which seems to indicate that it is caused just by the training and test division of the data.

Figures 5.8 and 5.9 may help better understand the results obtained in the hindcasting

Table 5.3: Results in terms of the RMSE for the field reconstruction (°C) obtained in ECA and ERA datasets (hindcasting problem), by the CRO-SL, HS, DE, Hill Climbing and greedy approaches.

	CRO-SL (ECA)			CRO-SL (ERA)		
	Best	Mean	Var	Best	Mean	Var
$N = 5$	1.74	1.74	0	1.54	1.55	0.0033
$N = 10$	1.59	1.61	0.0050	1.43	1.44	0.0076
$N = 15$	1.54	1.57	0.0051	1.40	1.45	0.0030
$N = 20$	1.51	1.52	0.0029	1.38	1.42	0.0034
	HS (ECA)			HS (ERA)		
$N = 5$	1.74	1.74	0	1.55	1.55	0.0002
$N = 10$	1.60	1.60	0.0016	1.44	1.44	0.0104
$N = 15$	1.55	1.58	0.0042	1.42	1.42	0.0031
$N = 20$	1.54	1.54	0.0021	1.41	1.41	0.0017
	DE (ECA)			DE (ERA)		
$N = 5$	1.77	1.77	0.0013	1.58	1.59	0.0097
$N = 10$	1.63	1.63	0.0034	1.47	1.48	0.0059
$N = 15$	1.57	1.58	0.0078	1.44	1.45	0.0069
$N = 20$	1.55	1.55	0.0053	1.43	1.49	0.0953
	Hill Climbing (ECA)			Hill Climbing (ERA)		
$N = 5$	1.81	2.11	0.0419	1.64	1.96	0.0630
$N = 10$	1.6540	1.82	0.0090	1.51	1.70	0.0109
$N = 15$	1.61	1.72	0.0043	1.48	1.60	0.0057
$N = 20$	1.57	1.66	0.0024	1.45	1.55	0.0031
	Greedy (ECA)			Greedy (ERA)		
$N = 5$	2.20	-	-	1.82	-	-
$N = 10$	2.10	-	-	1.72	-	-
$N = 15$	2.09	-	-	1.73	-	-
$N = 20$	2.09	-	-	1.68	-	-

problem. As can be seen, results in terms of location of the representative measuring points are quite similar to the previous partition of the datasets. In the ECA dataset, the solution with $N = 5$ is the same than in the previous case, and in the larger cases, the solutions found are not exactly the same, but they are very close. This indicates a consistency in the methodology applied, and a good performance of the CRO-SL. In the ERA dataset, the solution of the direct and hindcasting partitions are not the same, but again both solutions are very similar in structure, and the representative nodes selected are quite close in both problems.

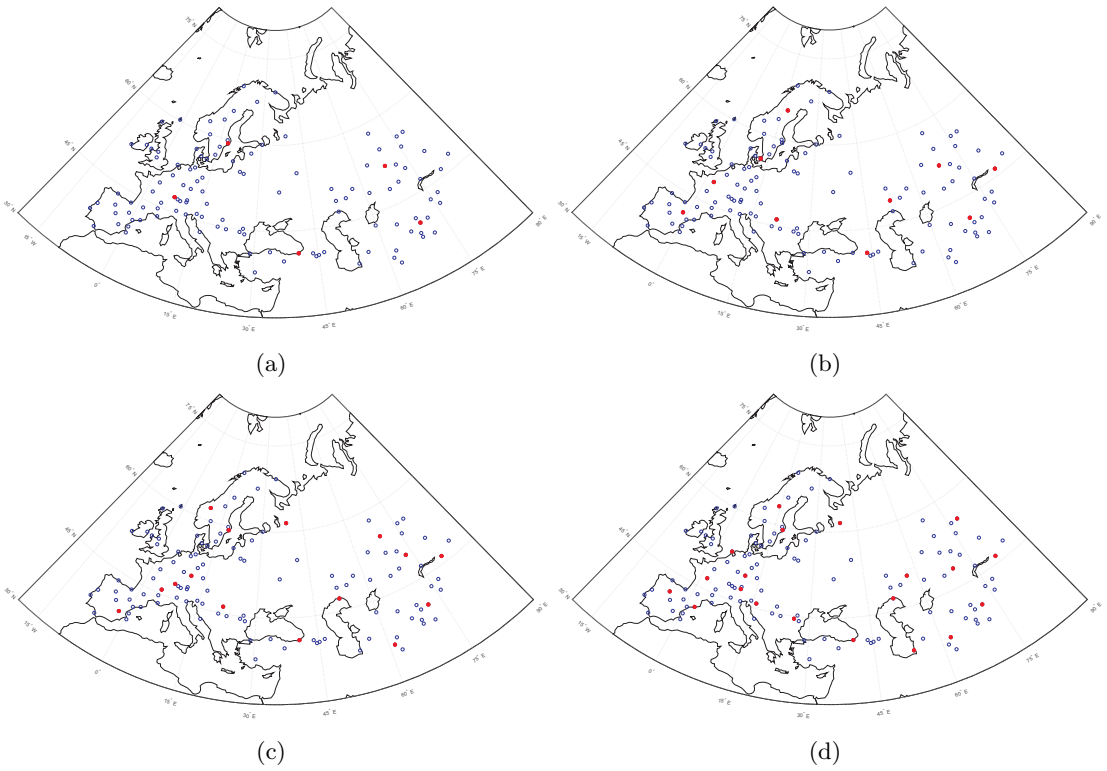


Figure 5.8: Best solution found (red points) by the CRO-SL, for the ECA dataset (hindcasting problem); (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

The consistency of the CRO-SL in the RS problem tackled in this paper can be further analyzed by trying to obtain the worst set of representative measuring stations/nodes with this algorithm. This can be done by inter-changing the minimization of the RMSE (Equation (1.6)) by a maximization of this measure. This way the solution obtained represents those measuring points with less information for the reconstruction algorithm (the AM in this case). Figure 5.10 shows an example of this for the case $N = 10$ in both datasets. Note how in both cases, the worst solutions group most of the selected measuring points at the top left-hand part of the considered studied area, and leave other zones uncovered, resulting in a lack of information for the AM reconstruction method. The RMSE for the ECA solution is 3.27 °C, whereas the RMSE for the ERA case is 4.59 °C.

5.3.5 Discussion II: details on temperature fields reconstructions obtained

Next, we analyze the temperature field reconstruction obtained from the best solution found by the CRO-SL algorithm. In this case, we exemplify with the case $N = 10$, both for the ECA and ERA datasets. For both datasets we have available the reconstruction of the temperature

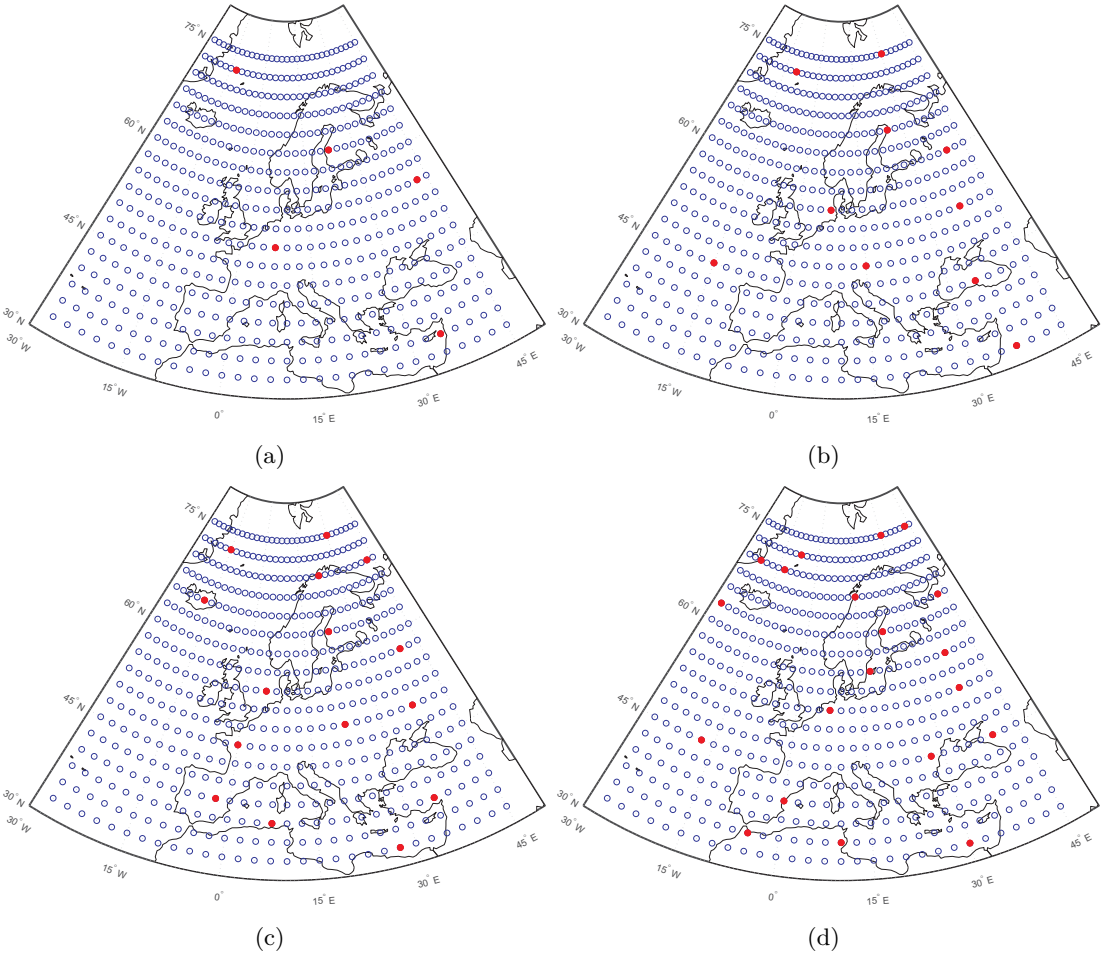
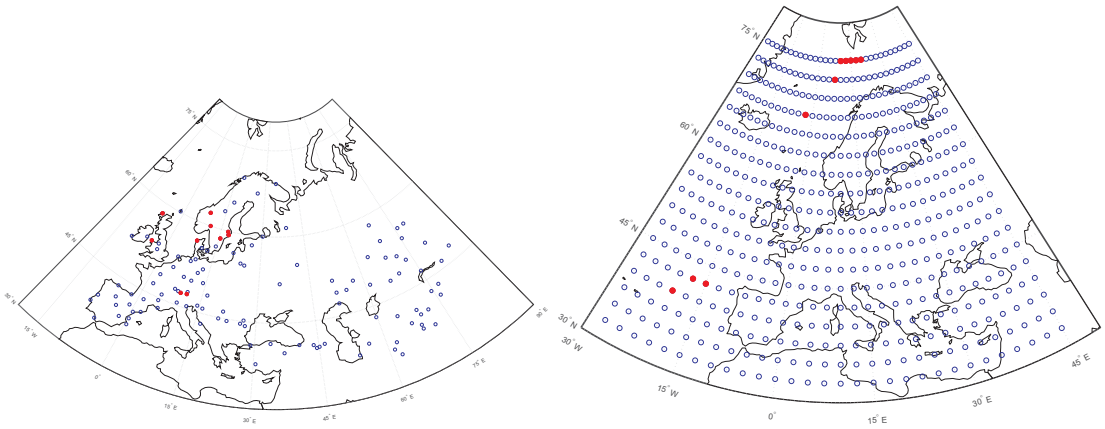


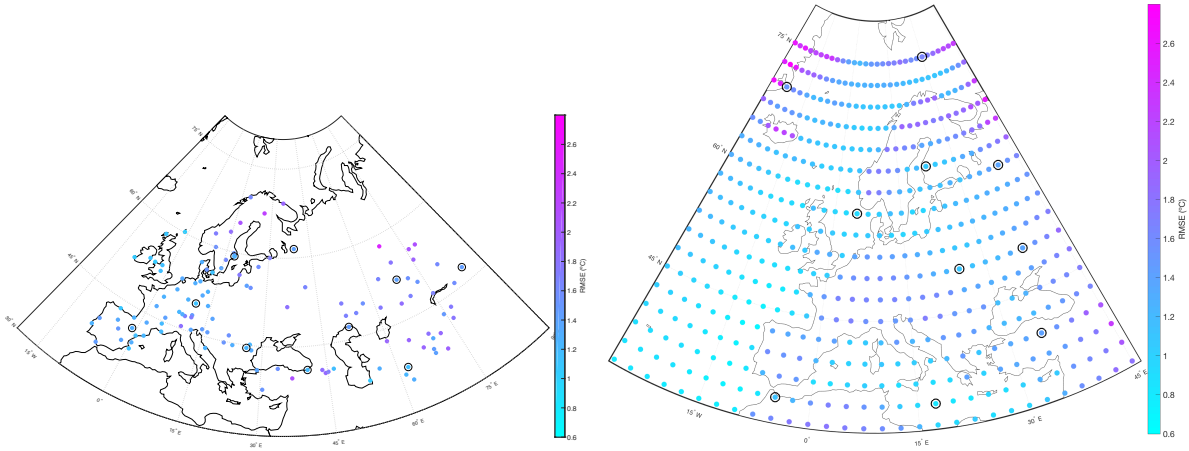
Figure 5.9: Best solution found (red points) by the CRO-SL, for the ERA dataset (hindcasting problem); (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

time series (in the test period), for all the stations/nodes, i.e. the reconstruction of 123 stations in the case of the ECA dataset, and 540 for the ERA. Figure 5.11 shows the average reconstruction error (whole test period) in all the stations of the data sets, for ECA (a) and ERA (b) datasets, with indication of the representative measuring points selected in each case. Note that the total reconstruction error is very good in general. In the ERA dataset it is possible to see that the measuring points over the ocean are better reconstructed than inland nodes, as could be expected. Furthermore, the distribution of selected nodes should be interpreted as a whole subset and not as individual points. This was evidenced when the CRO-SL outperformed the solutions obtained by the greedy algorithm (individualistic approach). We next show an example in the station/node with the best reconstruction error obtained. Figure 5.12 shows the reconstruction error during the whole test period in ECA (a) and ERA (b) datasets. It is possible to see how



(a) (b)

Figure 5.10: Set of least representative stations (red points) as inferred by the CRO-SL for $N = 10$ in ECA (a) and ERA (b) datasets.



(a) (b)

Figure 5.11: Reconstruction RMSE °C per station/node in the test period ($e(s_k, \mathbf{s}_{10}) = \sqrt{\frac{1}{t^V}\sum_{T=1}^{t^V}(F(s_k, T^*) - F(s_k, T))^2}$), in ECA and ERA datasets ($N = 10$ case).

the reconstruction error is significantly better in the ERA dataset, which is expected, since there is more information than in the ECA dataset (540 nodes versus 123 stations), and in addition the study area is smaller and the nodes are located on a regular grid. The reconstruction error in the ERA dataset varies between -1 and 1 °C in the majority of times reconstructed, with isolated peaks about 2 °C in the worst events. The reconstruction error in the ECA case moves mostly between -2 °C and 2 °C in the majority of times reconstructed, with peaks up to 4 °C at

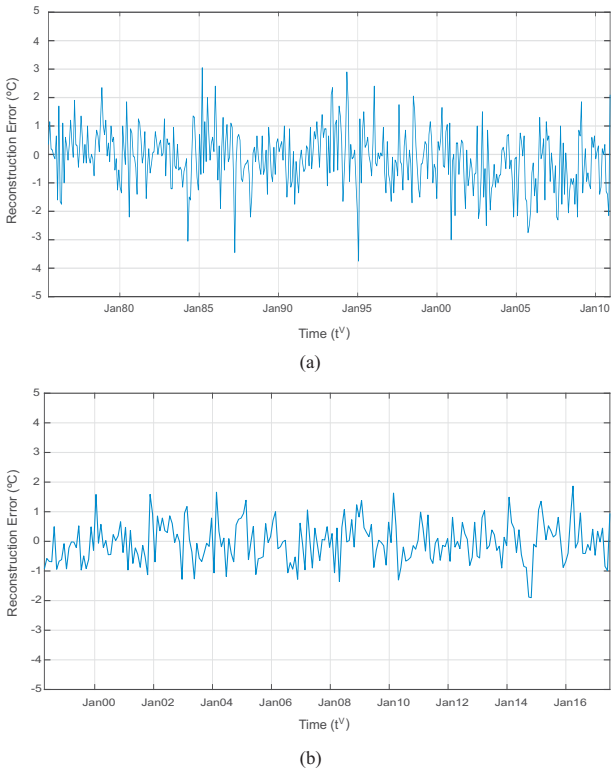
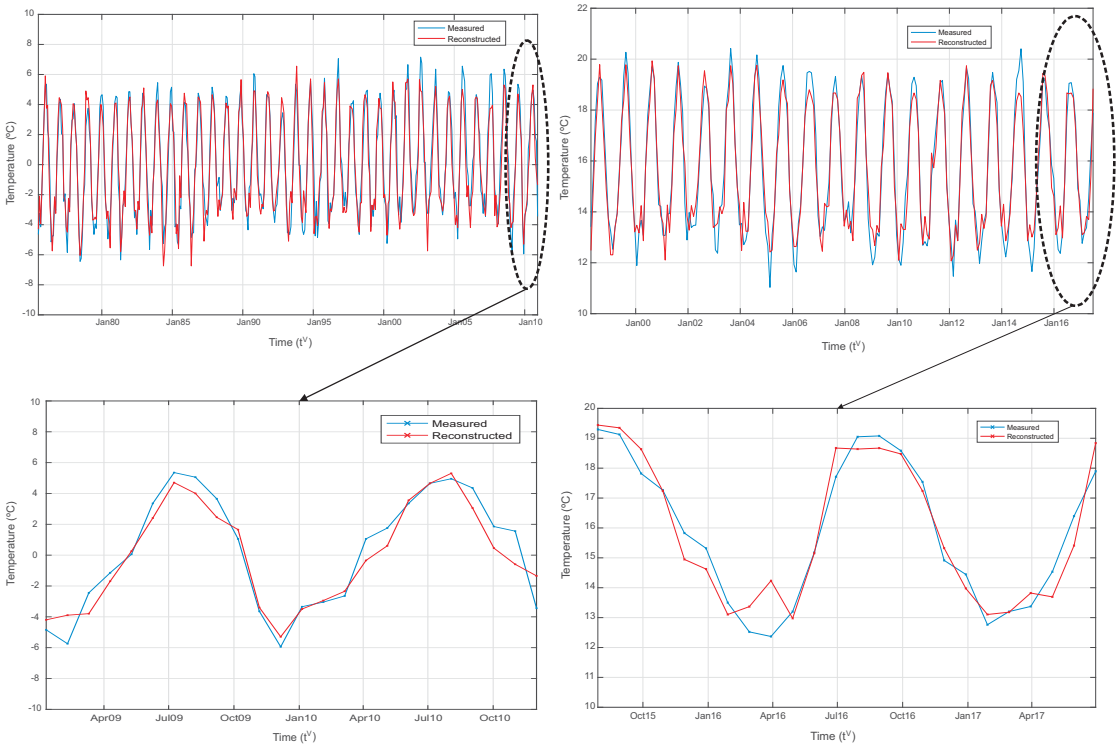


Figure 5.12: Reconstruction error $(F(s^*, T^*) - F(s^*, t))$ in ECA and ERA datasets (best station (s^*), $N = 10$ case); (a) ECA dataset; (b) ERA dataset.

some specific times. Figures 5.13 (a) and (b) show the temperature reconstructed for all the test period, and a detail on the reconstruction for the two final years of the test period, in the ECA and ERA best station, respectively. It is easy to see how the reconstruction of the temperature in both cases is accurate. The errors made can be better seen in the zoom for the final two years. In general, the temperature field reconstruction from the AM, based on $N = 10$ representative stations obtained with the CRO-SL (example displayed) is very good and robust in all the test period.

5.3.6 Discussion III: Consistency of the reconstructions from a climate perspective

On the climatic side, it is essential to assess the validity of the algorithm by addressing the consistency of the results at local and regional scales. For example, considering Figure 5.6 it is noteworthy to mention that, for a reduced subset of ECA data points ($N = 5$), optimal solutions tend to maximize the spatial coverage over the European continent. The best subset comprises points that are conveniently placed over five climatologically relevant regions: Central Europe,



(a) (b)

Figure 5.13: Real and reconstructed (AM method) temperature in the best station ($N = 10$ case) for the ECA abd ERA dataset (complete test period and zoom in the last two years); (a) ECA dataset; (b) ERA dataset.

Southern Scandinavia, Central Asia, Southern Asia and the Black Sea. All of them representing temperate and continental climates uniformly characterized by warm summers [Kottek2006]. Moreover, although the ERA reanalysis is based on a different set of temperature measurements, similar distributions are retrieved when the methodology is undertaken for the same number of stations (Figure 5.7). The resemblance of these distributions with independence of the input data, strengthens the climate consistency and spatial coherence of the obtained solutions. This climate concordance is preserved even for high values of N. In fact, note that, after a certain number of stations is reached, the improvement of the reconstructions is attained by increasing the density of data points in the surrounding areas nearby previously-defined zones, rather than by including new time series from distant regions. Furthermore, the consistency of the algorithm has also been addressed by applying the methodology to find the unsuitable distribution that leads to the worst climate reconstruction possible. Interestingly, these least representing nodes in Figure 5.10 are located in areas where they might be influenced by regional climate dynamics, such as the eastern Subpolar Gyre in the North Atlantic, and the amplification of sea surface

temperatures in the Nordic Seas, yielding temperature anomalies significantly different than the regional mean temperature [Sgubin2017, Lozier2008]. Additionally, the worst ECA subset is also conformed by points significantly influenced by regional climates. Such is the case of the nodes located in the Alps. Hence, climate field reconstructions carried out with these ill-favoured measuring points will not accurately represent the original climate history, and therefore the CRO-SL consistently identifies them as worst-case distributions. These meaningful findings show that the algorithm operates with climatic coherence, validating the applicability of this methodology.

5.3.7 Discussion IV: Results at different time scales

An interesting comparison can be carried out by analyzing the results obtained by the CRO-SL when considering different time scales in the data. For this, daily, monthly, quarterly and annual averages have been obtained from the ERA database. The idea is to compare the RS points obtained in these temperature fields with different time scales, and observe whether different patterns appear (or not) in the points selected by the CRO-SL. Experiments for $N = 5$ to $N = 20$ have been again considered in this case.

As can be seen in Figure 5.14, the daily and monthly cases present very similar patterns of RS points selected, with only minor differences. The quarterly scale case is also similar. The differences in the pattern are more accused in the annual average scale, where there are two points covering the Northern extreme zone of the Atlantic Ocean.

The same analysis with $N = 10$ is even more revealing. Figure 5.15 shows this case for the different scales considered. As can be seen in this figure, daily and monthly averages are quite similar, with slight variations in the points selected. However, quarterly and annual scales are different, and there is a clear trend to select points situated norther and over the ocean than in daily and monthly cases. This trend is specially significant at the annual scale, where 6 out of 10 points are situated over the Northern Atlantic ocean, whereas in the daily and monthly cases only 1 or 2 points cover this zone. A possible explanation for this result is that in the short-term scales (daily and monthly) the points covering continental zones provide more information. However, in the long-term scales, quarterly and annual, the filtering effect of long-term averages makes that points in the Ocean are more informative than those in-land.

Figures 5.16 and 5.17 show the results for $N = 15$ and $N = 20$ cases, respectively. It is possible to see a clear pattern, where the solution for the daily case is modified when longer-term averages are considered, and again, there are clear differences between the daily and monthly cases with the quarterly and annual, where more points over the ocean are considered. In all cases, the South West corner of the considered region is not covered with any point, which indicates that this zone is not very informative for the temperature field reconstruction in Europe. It might seem that the Azores High present at this zone is the responsible for this effect, because of the reduced variability in that area. It results in an information loss related to temperature field reconstruction from that specific zone.

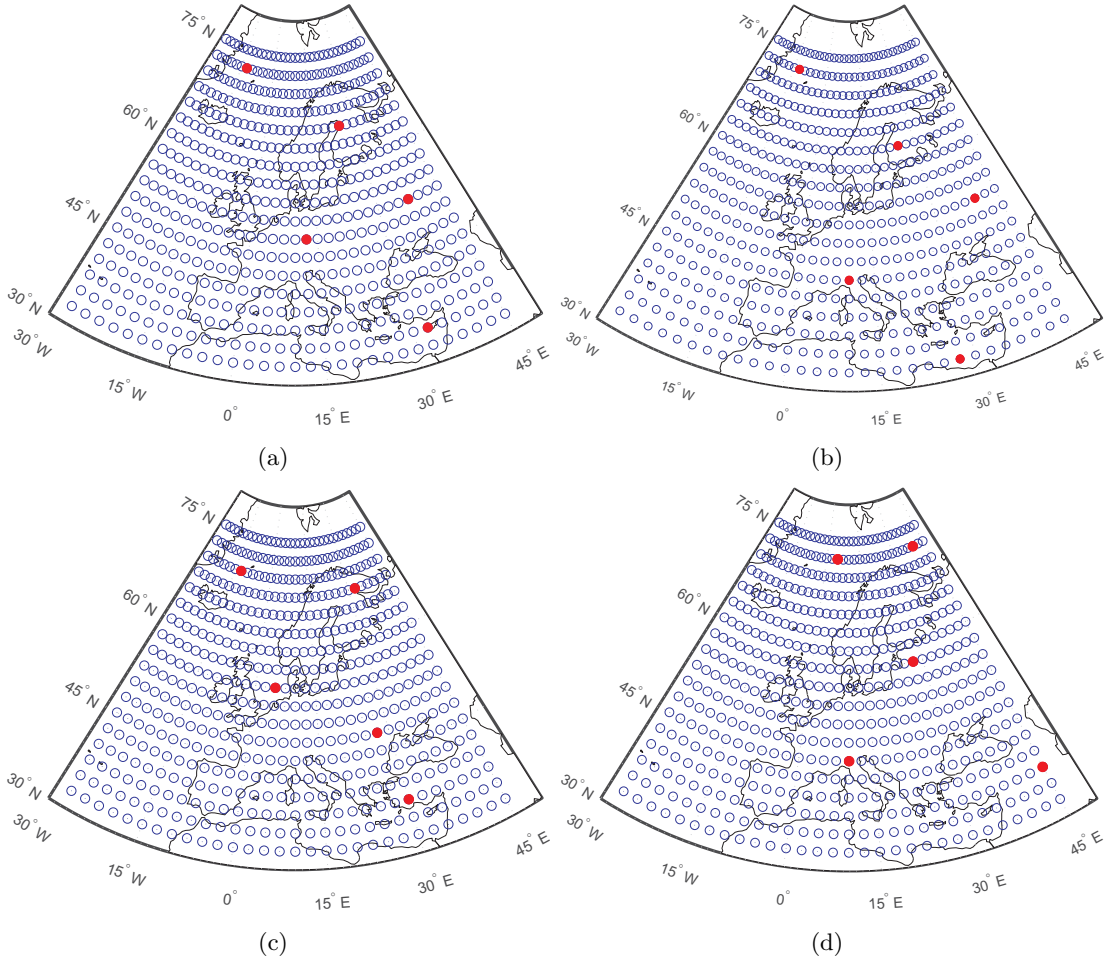


Figure 5.14: Best solution found (red points) by the CRO-SL, for the ERA dataset with different temporal scale ($N = 5$); (a) daily; (b) Monthly; (c) Quarterly and (d) Annual.

5.3.8 Discussion V: CRO-SL computational performance

Finally, a brief note on the computational behaviour of the CRO-SL algorithm. We discuss the CRO-SL performance in terms of how the different substrates (search procedures) help the algorithm find better solutions. Figure 5.18 shows the number of new larvae (solutions) able to get into the reef during the CRO-SL evolution, and the percentage of best larvae formed in each substrate, for the ECA and ERA datasets and $N = 10$. In this figure it is easy to see that the 2Px and the MPx operators dominate the production of the best larvae, and they are those which get the highest number of larvae into the reef in each generation. In the first stages of the algorithm, the contribution of the HS, DE and SM operators is significant, mainly the SM, which in the first 100 generations seems to produce better larvae than the MPx in the ECA dataset.

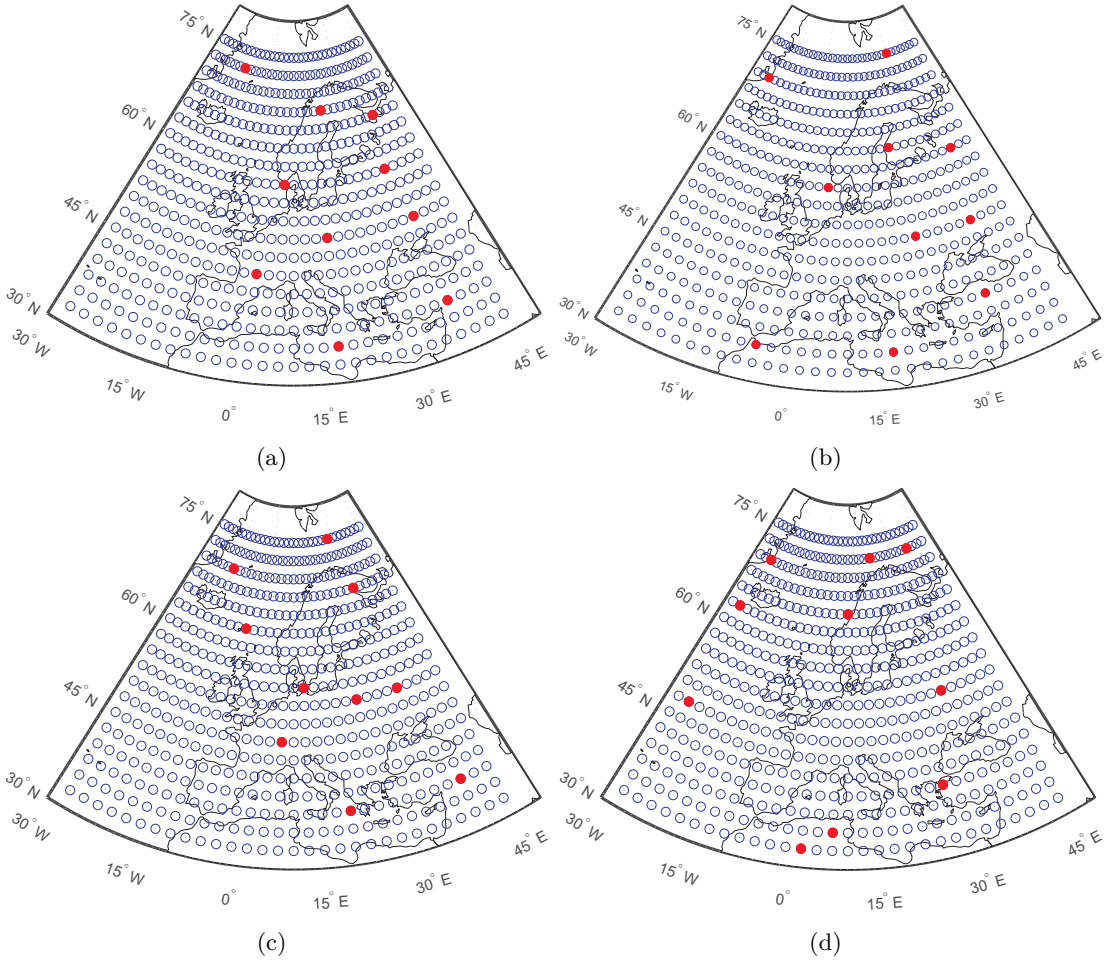


Figure 5.15: Best solution found (red points) by the CRO-SL, for the ERA dataset with different temporal scale ($N = 10$); (a) daily; (b) Monthly; (c) Quarterly and (d) Annual.

In the ERA case, the SM operator is the third best contributor to the search capability of the CRO-SL. Figures 5.18 (a) and (c) show a nice convergence of the algorithm in terms of number of new larvae in the reef versus generations. Note that at the beginning of the algorithm, the number of new larvae in the reef is significant, pushed by the fact that the number of holes in the reef is high, and the operators are able to find good quality solutions quickly. In the last stages of the algorithm, the number of new larvae in the reef is less important, and a minimum of larvae renewal due to depredation is maintained until the end of the algorithm. In the case of the best larvae generated, Figure 5.18 (b) and (d), note that there is a clear dominance of 2Px and MPx, consistent during all the CRO-SL evolution. Note also that, although these figures show the percentage of best solutions obtained per generation, not all them will be able to get attached in the reef. Thus, these figures only show the best larvae in generation, but not as part

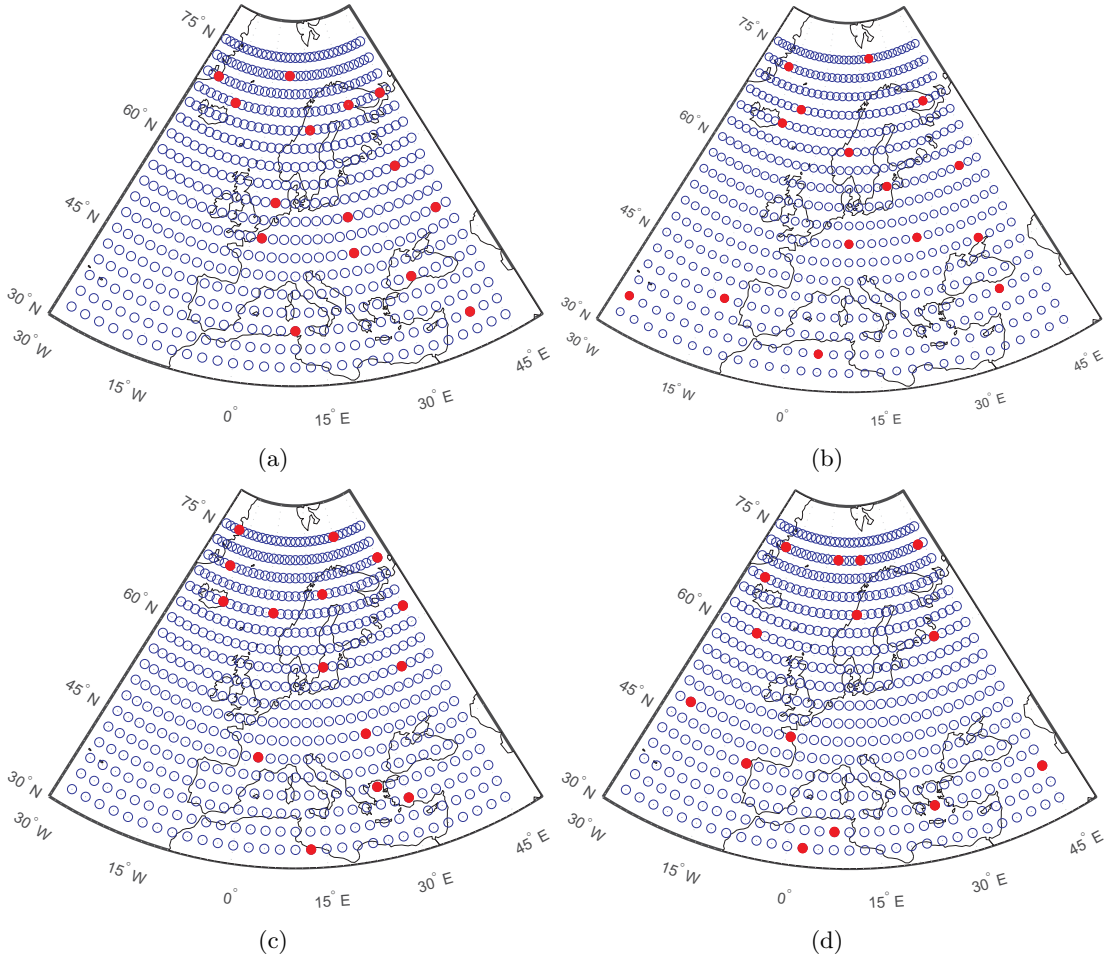


Figure 5.16: Best solution found (red points) by the CRO-SL, for the ERA dataset with different temporal scale ($N = 15$); (a) daily; (b) Monthly; (c) Quarterly and (d) Annual.

of the final reef of solutions considered in the CRO-SL (many of them will be discarded by the algorithm before getting into the reef).

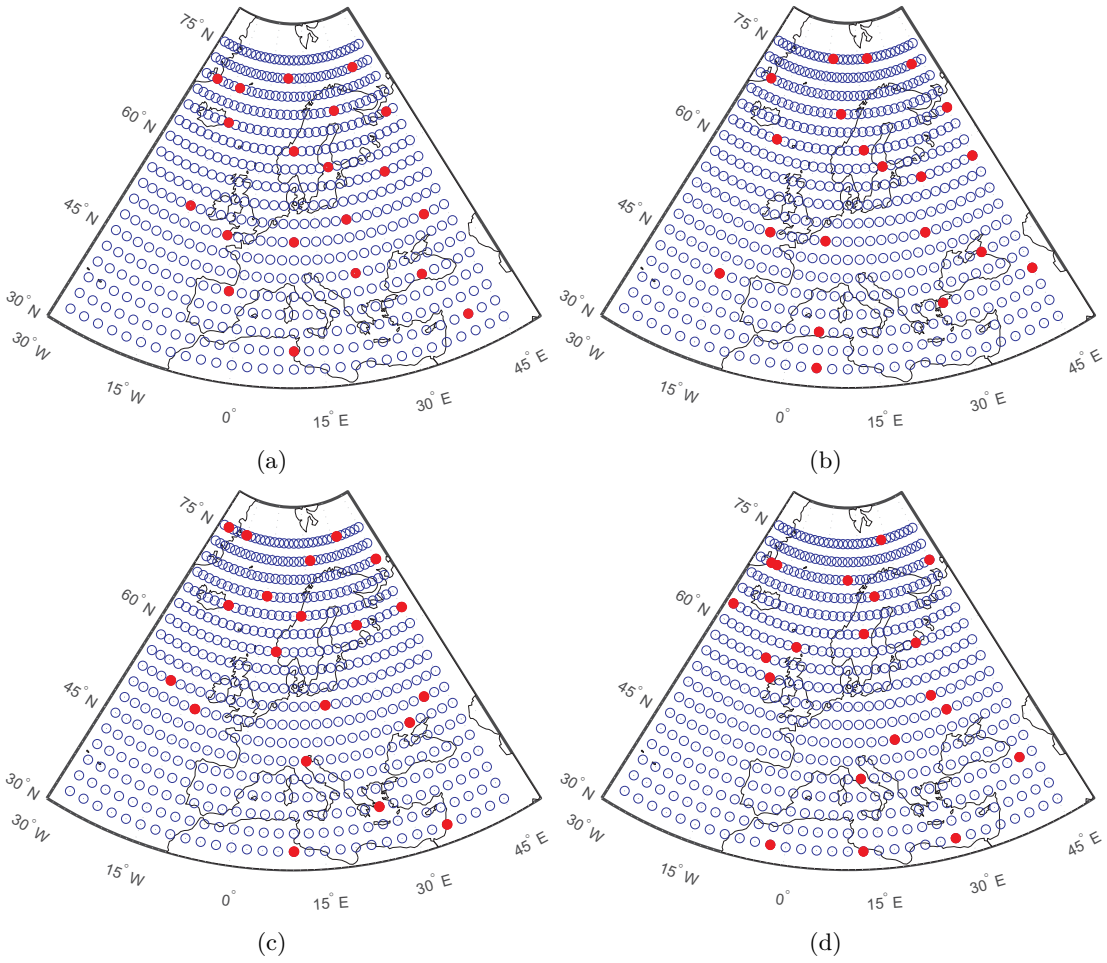


Figure 5.17: Best solution found (red points) by the CRO-SL, for the ERA dataset with different temporal scale ($N = 20$); (a) daily; (b) Monthly; (c) Quarterly and (d) Annual.

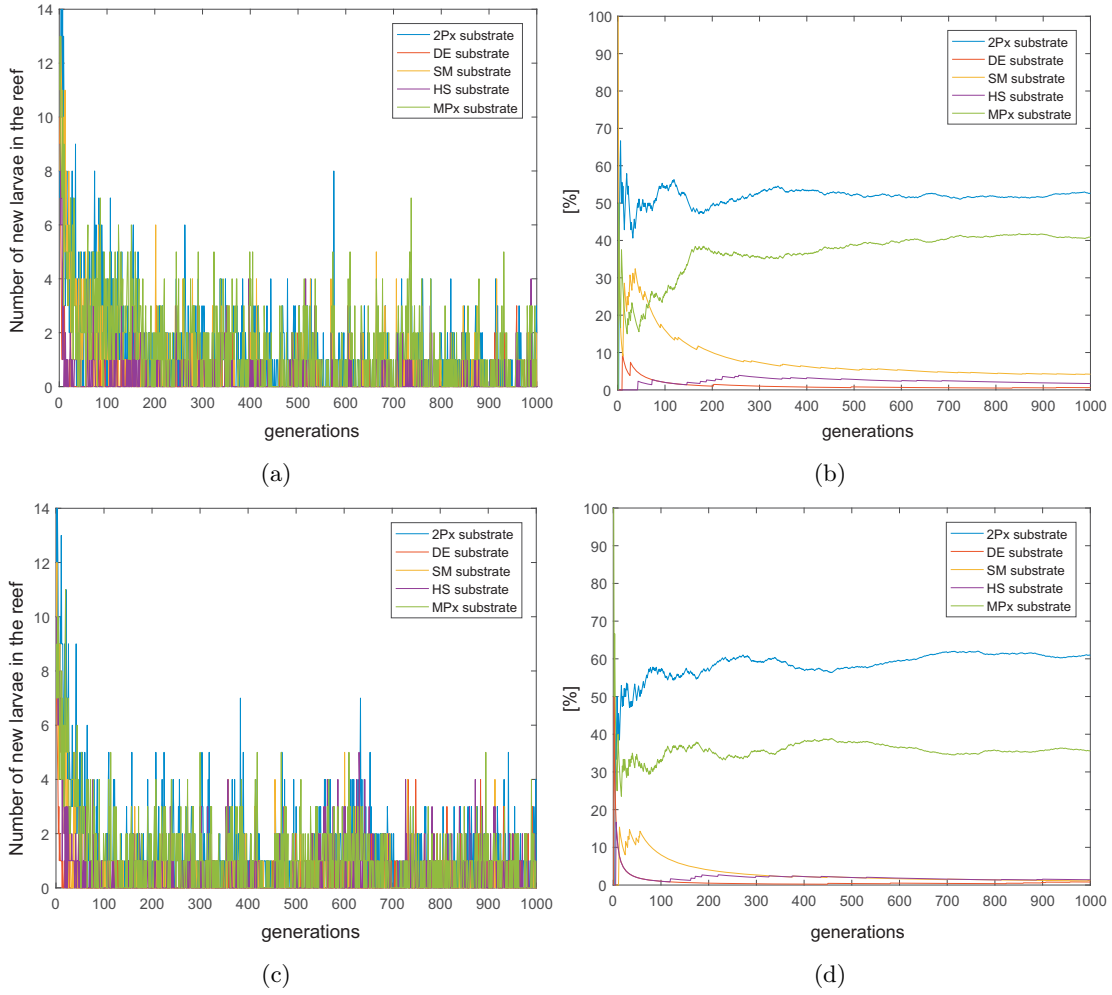


Figure 5.18: Evolution of the number of new larvae which are able to get into the reef per generation and substrate, and percentage of best larvae obtained from each substrate, for the ECA and ERA datasets in the case $N = 10$; (a) Number of new larvae in the reef (ECA); (b) Percentage of best larvae formed (ECA); (c) Number of new larvae in the reef (ERA); (d) Percentage of best larvae formed (ERA).

Chapter 6

Representative selection for wind speed fields reconstruction

6.1 Introduction

This chapter applies the methodology described in Section 5 to the wind speed, which is a variable much more difficult to predict, due to its extreme spatio-temporal variability. This limits the geographical representativeness of the wind measuring points. On the other hand, the density of wind observation sites is much lower than those for other variables, such as temperature or precipitation. Thus, characterising wind variability out of the observational points is usually much less reliable than for any other variable. In this chapter we analyze how wind fields can be reconstructed from a limited number of representative points. This problem is significant for wind speed prediction and wind resource analysis, for different reasons: the reconstruction of wind speed series is an important topic in wind speed prediction which is usually carried out by using neighbour stations [Saavedra2013, Dias2018]. The study of long-term wind speed variability is also relevant for wind resource analysis and evaluation [Ritter2017]. In these applications, the selection of a representative subset of points to characterize the wind speed is a challenging problem. In fact, traditional optimization algorithms do no work properly due to the discrete nature of the encoding, and the black-box characteristic of the objective function (outcome of the AM reconstruction algorithm). From a different point of view, the solution of this problem should allow the identification of the best location of measuring points to accurately evaluate a given wind field. As in the previous chapter, we deal with a RS optimization problem, in this case with a wind speed field. The CRO-SL algorithm, with the same structure, encoding and substrates than those defined in Section 5.2.2, has been considered.

6.2 Experimental evaluation

6.2.1 Data, algorithms for comparison and experimental parameters

In order to illustrate the performance of the chosen approach, we consider gridded monthly average wind data from the ERA-Interim reanalysis of the ECMWF [Dee2011]. Data from January 1979 until July 2017 are taken, with a total of $\hat{t} = 462$ months available. We have divided them into training period ($t^T = 231$) months and test period ($t^V = 231$) months. Figure 6.1 shows the location of the reanalysis nodes ($|S| = 540$ nodes considered).

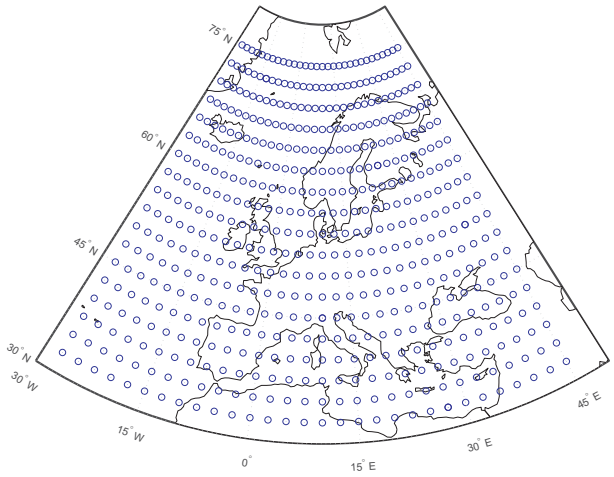


Figure 6.1: Location of the measuring points in the ERA-Interim reanalysis nodes considered.

For comparison purposes, we consider a number of state-of-the-art meta-heuristic algorithms: a HS [Geem2001], an EA [Eiben2003] with two points crossover, tournament selection and uniform mutation, an integer version of the PSO [Kennedy1995] and a SA [Kirpatrick1983]. The parameters of the CRO-SL used in this application are summarized in Table 6.1. This table also shows the parameters of alternative approaches for comparison.

6.2.2 Results and discussion

The experiments consist of running the CRO-SL and the alternative meta-heuristics for comparison in four different cases: $N = 5, 10, 15$ and 20. Note that they imply a reduction in the number of measuring points of 99%, 98%, 97% and 96%, respectively, to carry out the wind field reconstruction. Table 6.2 shows the results in terms of the RMSE in the wind speed field reconstruction, with the CRO-SL and the meta-heuristics considered for comparison. As can be seen, the CRO-SL is able to obtain the best results consistently for all values of N. The EA obtains the second best solution in most cases, slightly better than HS, PSO and SA, respectively. The differences in performance are more acute when comparing the CRO-SL with the rest of meta-heuristics, which means that the search capabilities of the proposed approach

are significant.

Table 6.1: Parameters of the optimization meta-heuristics compared in this paper: CRO-SL, HS, EA, PSO and SA.

CRO-SL	Parameters
Initialization	Reef size = 50×40 (2,000 positions), $\rho_0 = 0.9$
External sexual reproduction	$F_b = 0.80$, $\mathcal{T} = 5$ substrates: HS, DE, 2Px, MPx, SM
Internal sexual reproduction	$1 - Fb = 0.20$
Larvae setting	$\kappa = 3$
Asexual reproduction	$F_a = 0.05$
Depredation	$F_d = 0.15$, $P_d = 0.05$
Stop criterion	$k_{max} = 1000$ iterations
HS	**Parameters**
Harmony Memory	2,000 harmonies
HMCR	0.8
PAR	0.3
Stop criterion	$k_{max} = 1000$ iterations
EA	**Parameters**
Population	2,000 individuals
Selection	Tournament
Crossover	Two-points
Mutation	Uniform integer
Stop criterion	$k_{max} = 1000$ iterations
PSO	**Parameters**
Swarm	2,000 particles
Learning rates	$\phi_1 = 0.5$, $\phi_2 = 0.5$
Maximum velocity	$V_{max} = 10$
Stop criterion	$k_{max} = 1000$ iterations
SA	**Parameters**
Temperature Change	$T_{n+1} = \frac{T_n}{0.01 \cdot (n+1)}$
Changes per temperature	2,000
Mutation	Standard integer Mutation (SM)
Stop criterion	$k_{max} = 1000$ iterations

Once we have spotted the CRO-SL as the best meta-heuristics among those tested, we apply it in order to study the RMSE dependence of N for large number of N values. The results are shown in Figure 6.2. Note that each point corresponds to 5 CRO-SL runs (average value is displayed). Results for $N = 5, 10, 15$ and 20 are marked in red in the figure, for reference. As can be seen, the RMSE obtained is smaller when N grows, as expected, until a value of around

Table 6.2: Results in terms of the RMSE for the wind speed field reconstruction (m/s) obtained in the Europe reanalysis data, by the CRO-SL, HS, EA, PSO and SA approaches.

	CRO-SL		
	Best	Mean	Var
$N = 5$	**0.76**	0.77	$3e^{-2}$
$N = 10$	**0.71**	0.72	$2e^{-2}$
$N = 15$	**0.68**	0.69	e^{-2}
$N = 20$	**0.67**	0.68	$2e^{-2}$
	HS		
	Best	Mean	Var
$N = 5$	0.80	0.85	e^{-2}
$N = 10$	0.76	0.80	$2e^{-2}$
$N = 15$	0.74	0.77	e^{-2}
$N = 20$	0.72	0.76	$3e^{-2}$
	EA		
	Best	Mean	Var
$N = 5$	0.77	0.78	$3e^{-3}$
$N = 10$	0.73	0.74	e^{-3}
$N = 15$	0.72	0.73	e^{-3}
$N = 20$	0.69	0.71	$2e^{-3}$
	PSO		
	Best	Mean	Var
$N = 5$	0.78	0.81	$3e^{-3}$
$N = 10$	0.74	0.78	$2e^{-3}$
$N = 15$	0.73	0.77	$2e^{-3}$
$N = 20$	0.71	0.74	e^{-3}
	SA		
	Best	Mean	Var
$N = 5$	0.79	0.80	$6e^{-3}$
$N = 10$	0.76	0.77	$3e^{-3}$
$N = 15$	0.74	0.76	$2e^{-3}$
$N = 20$	0.72	0.73	e^{-3}

$N = 70$ stations, where it remains more or less constant. Note that the RMSE starts growing again for high values of N (from $N = 350$ stations), where it seems that the information is somehow degraded by the noise introduced when including a high number of measuring points

in the wind field reconstruction.

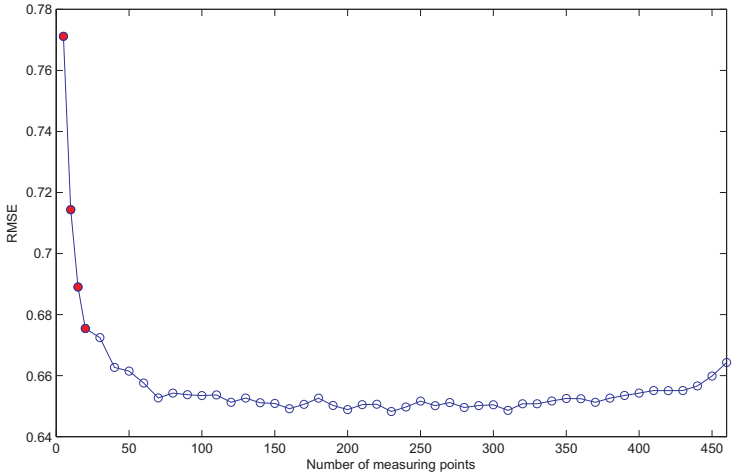


Figure 6.2: RMSE (m/s) obtained with the CRO-SL approach for different values of N.

Figure 6.3 shows the best solutions obtained by the CRO-SL (points selected), for the different cases considered. It is very interesting that the algorithm consistently selects points in the North Atlantic as those most informative for reconstructing the field of monthly wind speed, mainly in the cases with lower N. When the number of possible measuring points grows ($N = 15$ and 20), the algorithm includes points over the Atlantic ocean, and additional a few points in the Mediterranean, as those with the most information to reconstruct the wind speed field. This result is clearly explained by the prevailing Westerlies in Europe, which make that points over the Atlantic ocean contain the most information to reconstruct the average monthly wind over the whole continent. Note that the best solution obtained does not include any in-land point. Instead, for every solution, all the points are located over the ocean. This is again consistent, since there are no local influences introduced by orography and topography.

Figure 6.4 shows the reconstruction error of the complete wind speed field (normalized by the average wind speed of each point) with the AM, for different values of N. The worst reconstruction is found in the North West Atlantic (Greenland and Iceland), parts of the Middle East and some parts of the Mediterranean Sea. This is consistent no matter the value of N considered to guide the reconstruction with the AM. In general the reconstruction RMSE for the complete wind speed field is good, as can be see in Table 6.2, with values around 0.67 to 0.77 m/s when the solutions by the CRO-SL are considered.

Figure 6.5 shows the reconstruction error for the best point of the wind field reconstructed with the AM, in the test period. As can be seen, the error is small (less than 0.5 m/s in all cases), and there are not big differences with $N = 5$ or $N = 20$. This indicates that the reconstruction method is strong even from a small number of measuring points to start with, so it is robust when dealing with scarce information for the reconstruction. Note that when the number of

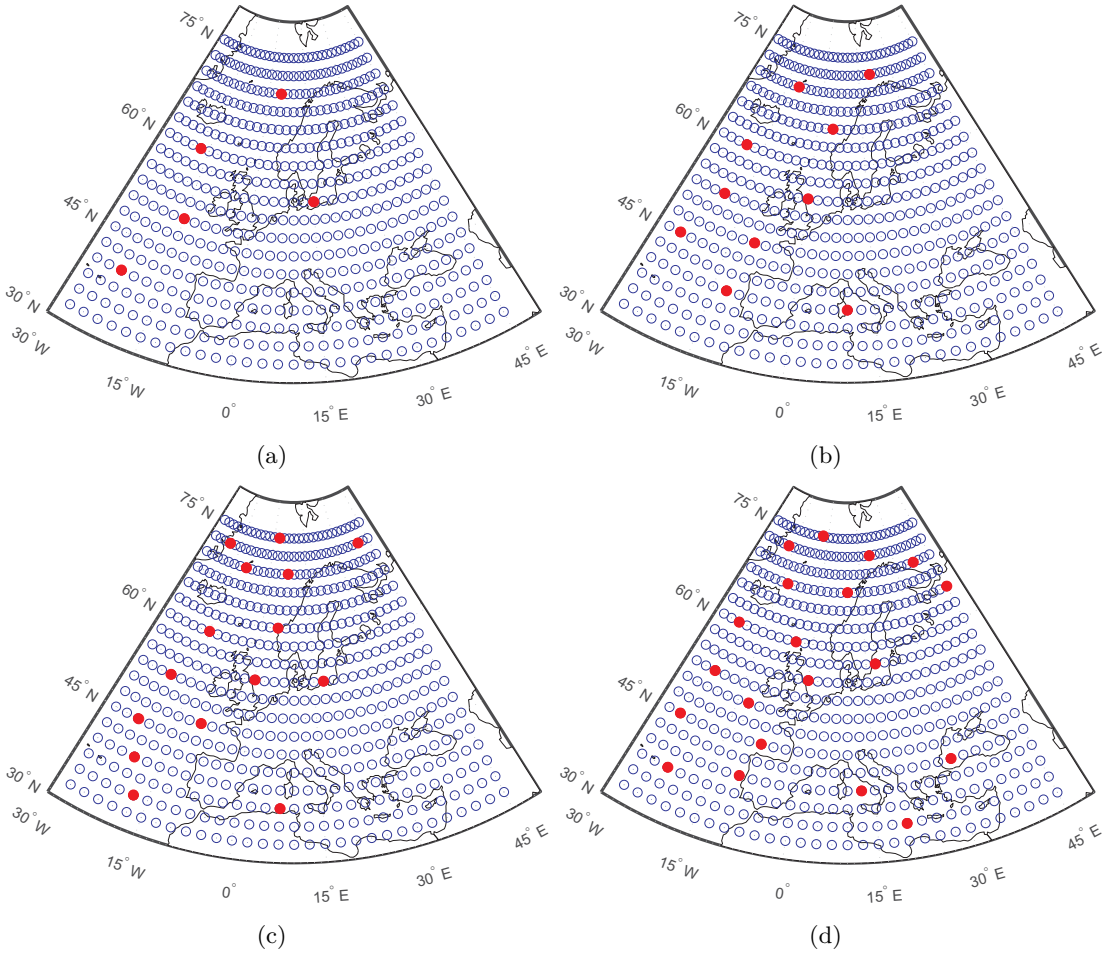
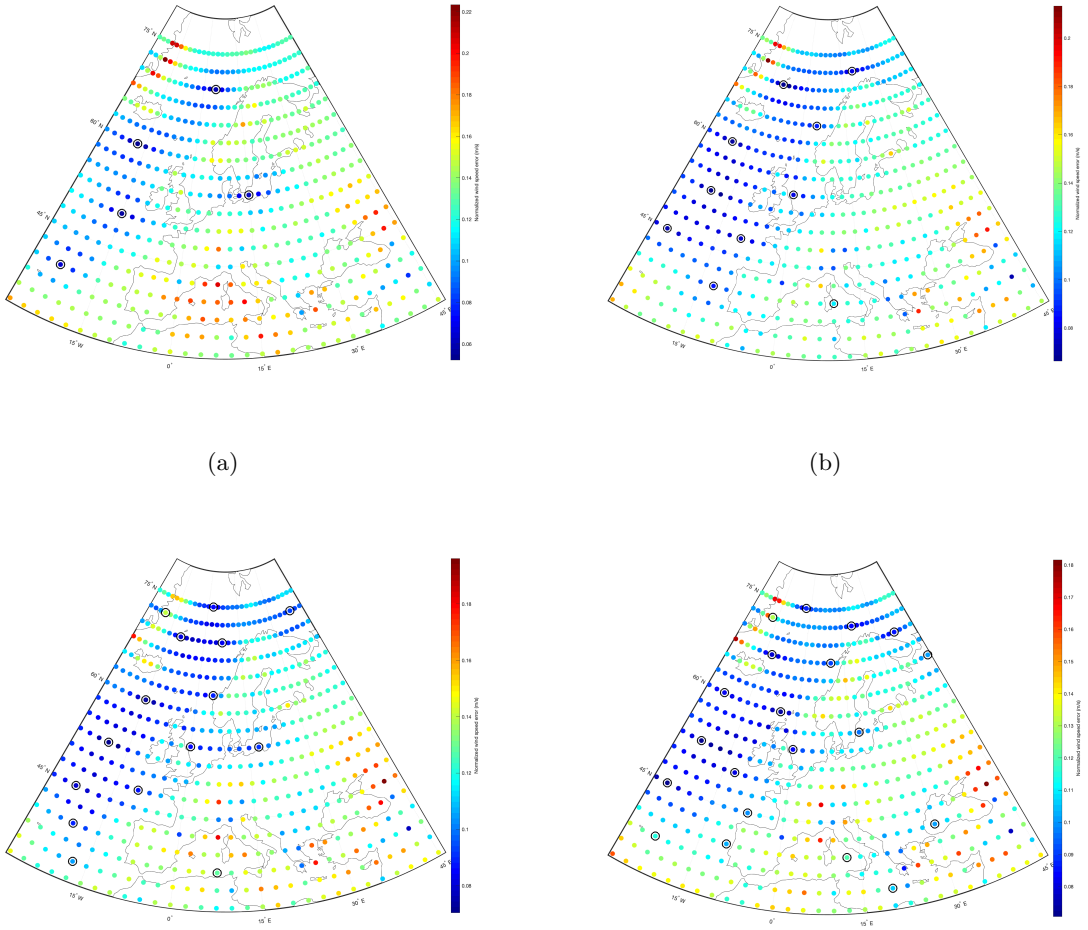


Figure 6.3: Best solution found by the CRO-SL (red points stand for the selected representative nodes); (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

representative points grows, the algorithm tends to better covering some specific zones (the Atlantic Ocean), instead of selecting new ones. This is significant, since aims to highlight the importance of the selected zones for the reconstruction, discarding the selection of points in alternative areas, in a similar way as for the temperature case.

Figures 6.6 (a), (b), (c) and (d) show the best solutions obtained by the CRO-SL in the monthly field reconstruction problem for the case $N = 20$, in different seasons: Spring, Summer, Autumn and Winter, respectively. It is possible to see how the representative points are again consistently located in the North Atlantic in all seasons. It seems that in Autumn and Winter the algorithm chooses some more points in the Mediterranean to characterize the wind speed field than in Spring or Summer. In general the seasonal study shows similar zones as those obtained when all the data are managed.



(a) (b)

(c) (d)

Figure 6.4: Reconstruction error of the wind speed field (normalized by the average wind speed of each point) with the Analogue method, for different number of selected representative points; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

Finally, we have identified the least representative points for the reconstruction (Figure 6.7). As can be seen, in all cases the selected points are mainly located at the North of Scandinavia and some points in the Eastern Mediterranean. The former are located close to circumpolar circulation, local winds with deficit of information in the reconstruction of the whole wind speed field. The latter could be related to the high seasonality of the circulation in that area, associated with the Etesian winds – local strong north winds of the Aegean Sea – which also seem to have

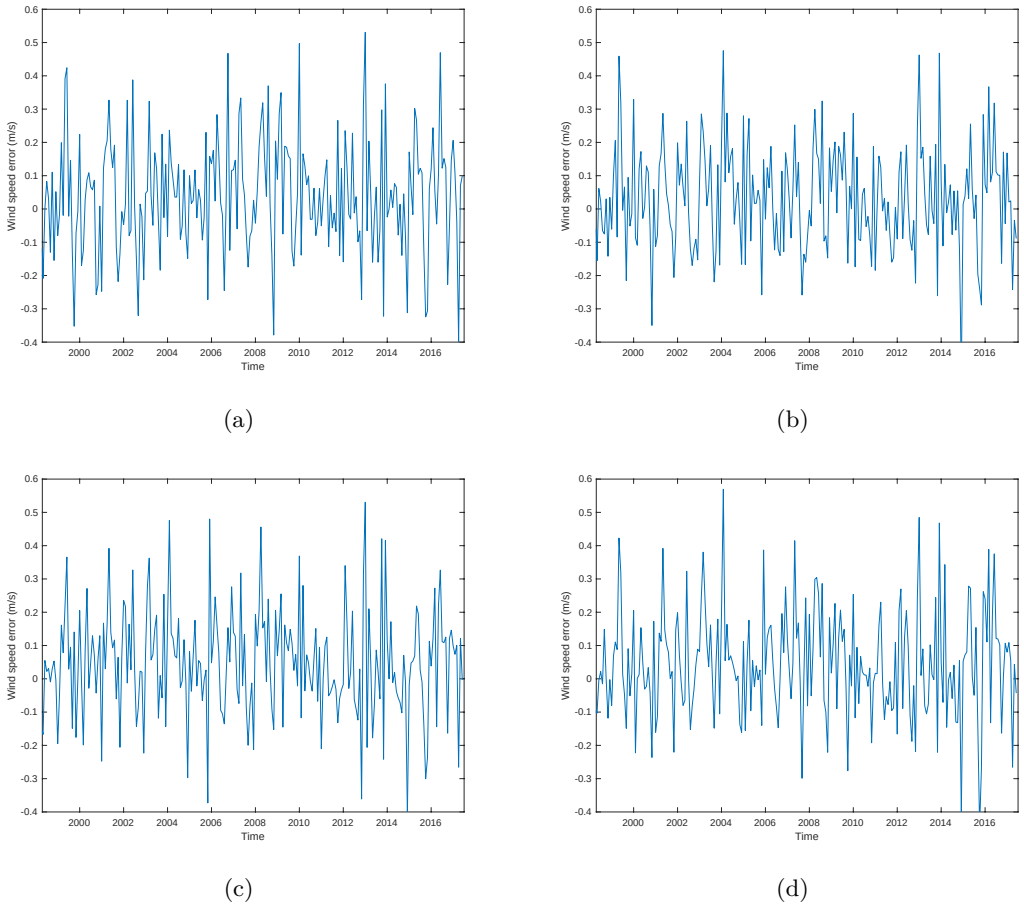


Figure 6.5: Reconstruction error for the best measuring point reconstructed with the Analogue method, for different number of selected representative points; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

little information for the reconstruction of the wind speed field in Europe. This analysis shows the physical consistency of the results obtained by the CRO-SL with AM also for wind. Note that the method does not consider any other climatological data source, but the wind speed series in the past, which reflects the robustness of the method.

Regarding the specific performance of the CRO-SL, we can study it in terms of how the different substrates (search procedures) help the algorithm find better solutions. Figure 6.8 shows the percentage of best larvae formed in each substrate, for the $N = 20$ case. It is easy to see that the 2Px and the MPx operators dominate the production of the best larvae, with the SM operator as a good contributor, specially in the first stages of the algorithm. The HS and DE substrates are those contributing the least to generate best larvae (solutions) in the algorithm's evolution. Another interesting analysis on the CRO-SL performance is given by

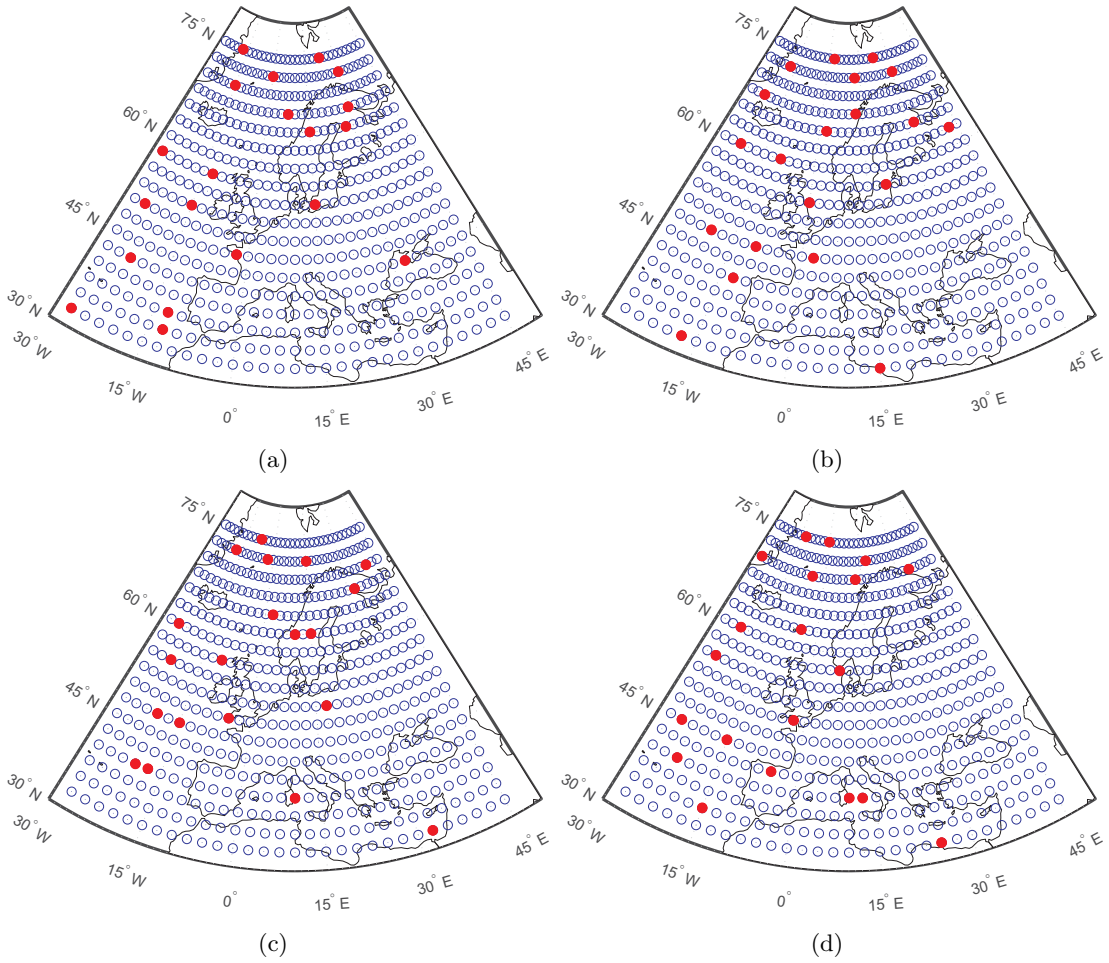


Figure 6.6: Best solution found by the CRO-SL (20 representative points), in different seasons; (a) Spring; (b) Summer (c) Autumn and (d) Winter.

Figure 6.9, which shows the number of new larvae able to get into the reef during the CRO-SL evolution. In this case, the contribution of the 2Px and SM operators is higher than the rest in the first stages of the algorithm, while the 2Px and MPx finally dominate the inclusion of new larvae in the middle and final generations of the CRO-SL. Note that at the beginning of the algorithm the number of new larvae in the reef is significant, pushed by the fact that the number of holes in the reef is high, and the operators are able to find good quality solutions quickly. In the last stages of the algorithm, the number of new larvae in the reef is less important, and a minimum of larvae renewal due to depredation is maintained until the end of the algorithm.

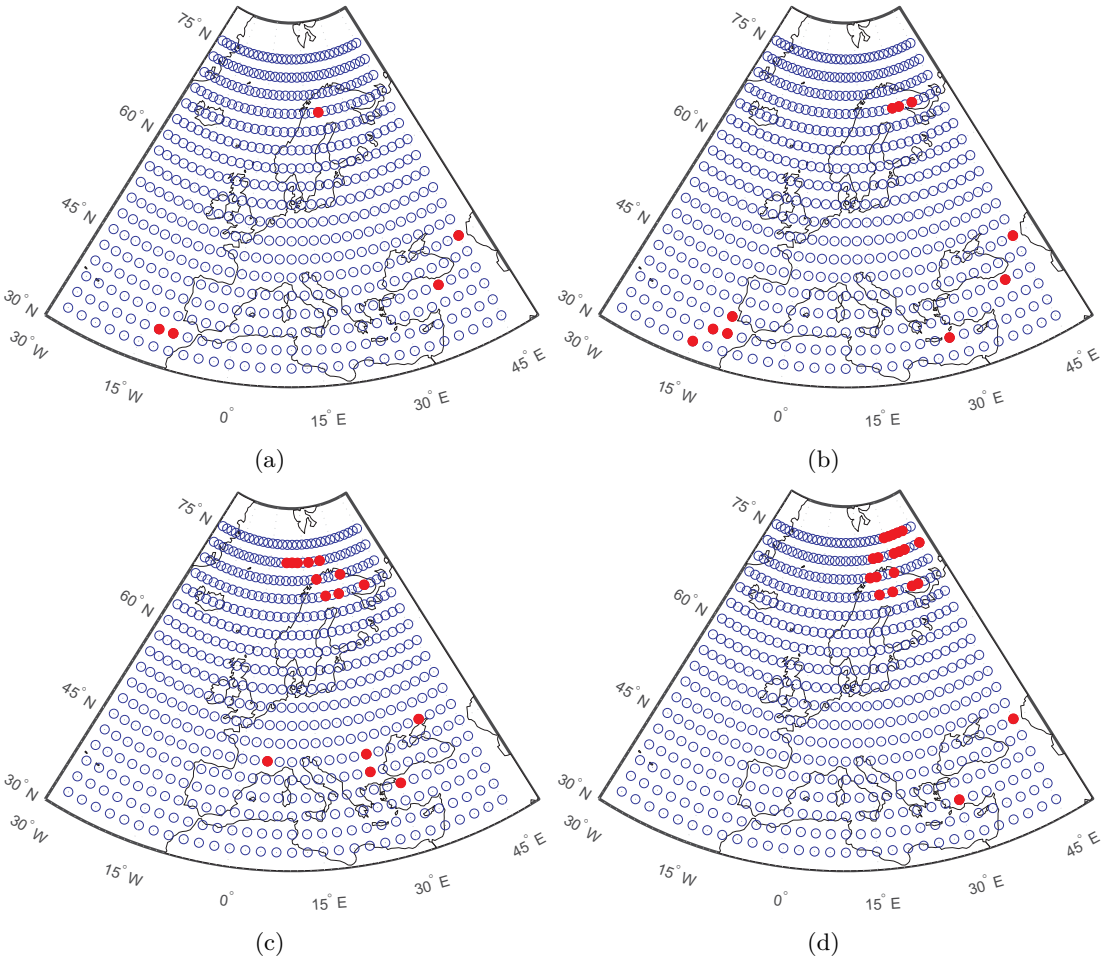


Figure 6.7: Least representative points for the wind speed field reconstruction; (a) $N = 5$; (b) $N = 10$; (c) $N = 15$ and (d) $N = 20$.

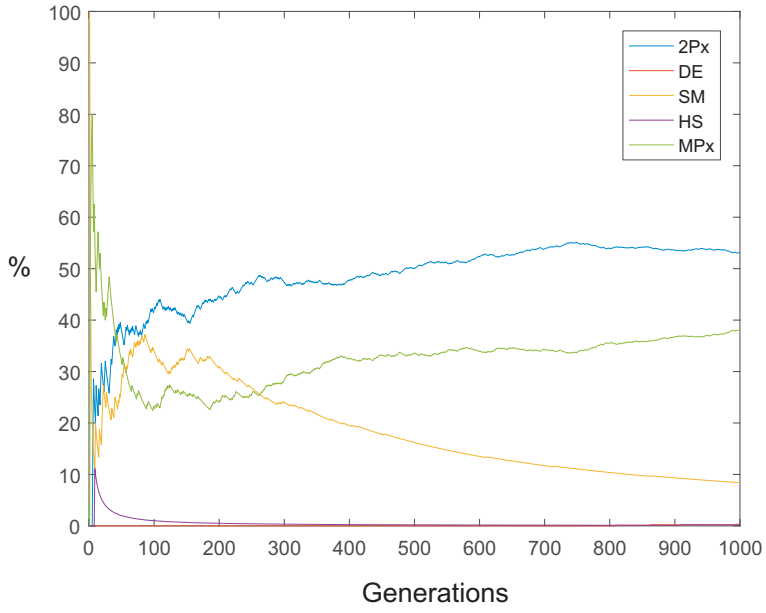


Figure 6.8: Percentage of best larvae obtained in the CRO-SL from each substrate, in the case $N = 20$.

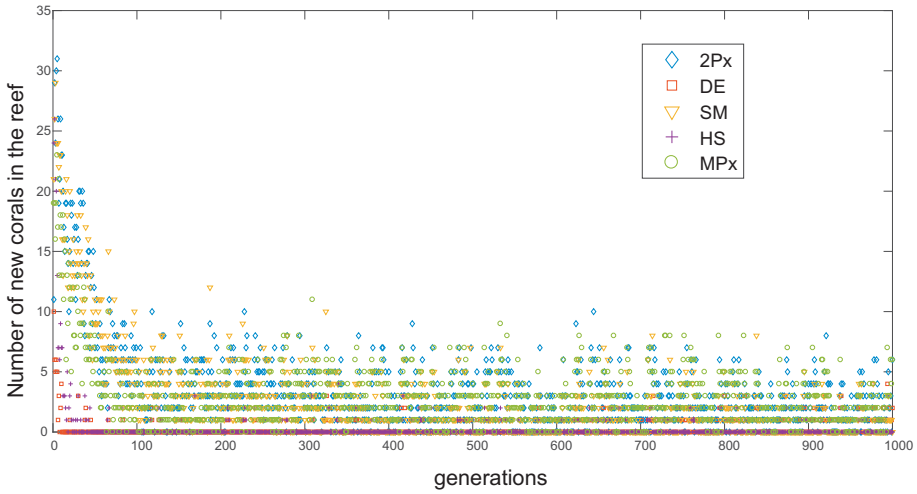


Figure 6.9: Evolution of the number of new larvae in the CRO-SL which are able to get into the reef per generation and substrate, case $N = 20$.

Chapter 7

Conclusions and future research

7.1 Conclusions

The research activity developed in this Ph.D. Thesis has produced significant contributions and results in two research areas: short-term forecasting problems of wind and solar energy resources and representativeness problems in fields of meteorological and climatological variables. Short-term forecasting of wind and solar radiation is key in the management of renewable energy facilities such as wind farms and solar photovoltaic plants. These renewable resources have problems of integration in the electric grid, due to their inherent intermittency, which may cause power transport problems if a good prediction of the resource is not available. Short-term forecasting of these renewable resources is then a cutting-edge problem, in which obtaining a high-quality forecast of wind and solar resource have important socio-economic repercussion. The second research area where this Thesis has produced research results is the representativeness problems in climatological and meteorological fields. It is well known that gridded fields of climatological and meteorological series have inherent complexity to be evaluated, mainly because the high number of variables involved in the process. This Thesis has obtained different procedures to reduce this complexity, focused on a reduction of points in the field with the minimum information loss. The evaluation of this procedure for complexity reduction has been carried out in temperature and wind data over Europe.

Regarding the computational part of the Thesis, we have used a recently proposed evolutionary-based optimization approach, known as Coral Reefs Optimization algorithm (CRO). This approach has been used to construct different hybrid Soft-Computing approaches, to face optimization problems related to prediction systems in Renewable Energy applications. The CRO and some improved versions of the algorithm have been mixed with Artificial Neural Networks such as Extreme Learning Machine approaches to solve problems of Feature Selection. The CRO has also been the core algorithm in the problems of Representative Selection points for optimal reconstruction of temperature and wind fields. A modified version of the algorithm, called CRO-SL has been used. In this case, the algorithm has been hybridized with the Analogue

Method to guide the search with physical sense.

A summary of the main contributions of the Ph.D. is given next, structured in two blocks: contributions on short-term prediction problems and contributions on representativeness problems in climatological and meteorological fields.

- **Short-term prediction problems of wind and solar energy resources:**

  1. We have evaluated two different novel FSP wrapper approaches based on the combination of the CRO and CRO-SL algorithms with an ELM. First, the basic CRO algorithm has been used to obtain the best set of variables in a problem of wind speed prediction in a wind farm located in the North-Western coast of the USA. An integer encoding is considered in this case in the CRO approach, so a fixed number of features are managed by the algorithm. The prediction capacity of each reduced set of features is analyzed by means of an ELM. A comparison with alternative approaches based on evolutionary computation has been carried out. The experiments have shown that the best result was obtained with a set of 9 features in the CRO, with a prediction error by the ELM of 2.5 m/s in the test set. Comparison with alternative regression approaches such as SVR have shown the effectiveness of using an ELM in the feature selection for this problem. In particular, a comparison with an EA has shown improvements up to 2% when the CRO is used as global search algorithm.
  2. In a second case of study the CRO-SL has been tested as global searcher, hybridized as well with a fast training ELM. We have evaluated this hybrid FSP algorithm with data from a Spanish wind farm. The robustness of this approach allows including different objective functions in the CRO-SL, in order to guide the FSP. With this, we have evaluated the possibility of improving the ELM prediction performance in different time-horizons (hourly and daily in this case). We have shown that is it possible to do so, with better improvement in the hourly time-horizon prediction. The comparison with alternative final regression has shown that the feature selection mechanism may affect their performance differently: we have analyzed the case of a SVR where the features obtained by the CRO-SL-ELM degrade the performance. On the contrary, the performance of an MLP seems to be enhanced by the feature selection pre-processing step. Thus, the wrapper approach performance may be very dependent on the final prediction mechanism applied.
  3. An important conclusion from both problems is that a fast-training prediction algorithm is needed for the feature selection step when the wrapper methodology is selected. Since a global search procedure is involved in this process, the faster the prediction algorithm training, the better the search carried out, since more generations or search cycles could be considered. In this sense, the ELM is really well fitted for this task, since it is extremely fast in its training sequence, and in addition its

prediction performance is very competitive. Alternative approaches related to filter feature selection mechanisms are known to offer a poorer performance than their wrapper FSP counterparts in prediction problems [Kohavi1997]. In particular cases with a huge number of very correlated features, hybrid filter-wrapper approaches may work, using a first step with a filter mechanism to remove the most related features [Huda2014, Huda2016]. Then a wrapper similar to those discussed in this chapter could be applied.

4. We have also solved a problem of GSR prediction using as predictive variables the outputs provided by a numerical weather model in a grid area over the target point (located at Toledo, Spain). The selection of the best predictive variables and the best grid points to be considered has been performed using the CRO-SP algorithm, and the prediction has been obtained using an ELM network. The ultimate goal has been to evaluate what predictive variables from the numerical weather model (i.e. the WRF model) perform the best. For this purpose, each species in the CRO-SP encodes a fixed and different number of variables to be analyzed and the best species comes out as a result of the co-evolution. To determine the best set and number of predictive variables, three experiments have been run in a 10-fold cross validation scheme and the RMSE has been used as common measure. The experiment where 7, 8, 9, 10 and 11 variables are co-evolved (experiment $\mathcal{E}_3$), produces an average error result of 69.19 W/m^2 and a lowest error of 68.03 W/m^2, turning in a 21.62 % and 22.03 % improvement, respectively, over the average and best prediction without feature selection. We show this way how a different version of the CRO approach (in this case considering species to manage different constraints of the problem) is able to provide very good results, improving the performance of the prediction system.

- **Representativeness problems in climatological and meteorological fields:**

1. We have first tackled a problem of selection of the optimal representative measuring points over a temperature field, out of a whole large amount of initial ones. It has connections with the more general RS problem in data science and can be stated as a hard discrete optimization problem. Here we deal with a specific version for robust reconstruction of monthly average temperature fields. We have used the CRO-SL algorithm together with the AM as the reconstruction algorithm. We have shown that the approach is able to deal with reconstruction problems in gridded and un-gridded datasets. In the experiments carried out in well-known temperature datasets in Europe and central Asia, the CRO-SL has been able to obtain solid solutions and robust reconstructions. Indeed, this method outperformed other problem-solving heuristics such as the greedy algorithm, leading to the conclusion that regional reconstructions cannot be well-represented by information retrieved from just a compilation of individual stations. We have also shown the consistency of the results by solutions of

a hindcasting problem, very similar to the initial direct problem. A final computational analysis on the best search procedures implemented in the CRO-SL for the RS problem has been carried out, showing the multi-point and two points crossovers are the most effective operators to obtain high quality solutions in this problem.

2. In the specific temperature fields databases considered for Europe and Centra Asia, the CRO-SL obtains solutions which maximize the coverage of the region, as could be expected from a RS point of view. In the ECA dataset, in the case of $N = 5$, the stations which are considered as the representatives cover central Europe, Scandinavia, Central Asia, Black Sea and Southern Asia. When $N = 10$, Europe is modelled with two representative stations, one for central Europe and another one for southern Europe located in Spain. The Scandinavian station is kept, and another close station is also selected in northern Russia. The Black Sea is in this case covered with two stations, one in the East and another one in the West, and the rest of Asia is covered by four stations, two of them in similar locations as in the $N = 5$ case. When N grows, the representative stations still cover the same areas, but in this case with increasing density, resulting in a more accurate temperature field reconstruction by applying the AM. In the ERA dataset, when $N = 5$ is considered, one node is located at the north of Iceland, to cover polar regions, another node is located in Scandinavia, a node in northern Italy covers central and southern Europe and the two remaining nodes cover eastern Europe and the Mediterranean with a node in Egypt. When $N = 10$, two nodes cover the polar region, three nodes are devoted to Scandinavia, two nodes cover central and eastern Europe, one node at the south of Spain covers southern Europe and there are other two nodes, for the Mediterranean and Black Sea regions. The solutions with larger N follow this trend, by locating more representative nodes in the pole and Scandinavia, selecting some of them for central and eastern Europe, and locating the rest in the Mediterranean and southern Atlantic zones. Regarding the accuracy of the fields reconstruction, the reconstruction error obtained in the ECA database varies between -2 °C and 2 °C in the majority of times reconstructed, with peaks up to 4 °C at some specific times. The reconstruction error in the ERA dataset varies between -1 and 1 °C in the majority of times reconstructed, with isolated peaks about 2 °C in the worst events.

3. An analysis of the RS results obtained by the CRO-SL when considering different time scales in the temperature data has been carried out. This analysis has been obtained only for the ERA database (regular grid points), where daily, monthly, quarterly and annual averages have been obtained and processed by applying the CRO-SL with the AM to extract the representative nodes for each case. The results obtained show that daily and monthly averages are quite similar, with slight variations in the representative points selected. However, quarterly and annual scales are different, and there is a clear trend to select points situated northwards and over the ocean than in daily

and monthly cases. This trend is specially significant at the annual scale, where 6 out of 10 points are situated over the Northern Atlantic ocean, whereas in the daily and monthly cases only 1 or 2 points cover this zone. A possible explanation for this result is that in the short-term scales (daily and monthly) the points covering continental zones provide more information. However, at the long-term scales, quarterly and annual, the filtering effect of long-term averages makes that points in the Ocean are more informative than those in-land. In all cases, the South West corner of the considered region is not covered with any point, which indicates that this zone is not very informative for the temperature field reconstruction in Europe. It might seem that the Azores High present at this zone is the responsible for this effect, because of the reduced variability in that area. This should result in an information loss related to temperature field reconstruction from that specific zone.

4. We have also tackled a RS problem related to the reconstruction of monthly averaged wind fields in Europe, from a reduced number of representative points ($N = 5, \ldots, 20$), also by applying the CRO-SL algorithm. As in the temperature field reconstruction case, the CRO-SL has been combined with the AM as reconstruction procedure. An experimental evaluation of the method using the ERA-Interim reanalysis data (monthly resolution) has been carried out. The results reveal that the spatial pattern of the representative points is consistently located over the Atlantic ocean, which agrees with the predominance of the mid-latitude Westerlies in the region, whereas in-land measuring points have less reconstructive information. The reconstruction of the complete wind speed field with a reduced number of selected points is accurate, with average errors lower than 1 m/s in the test period.

5. We have observed a common behaviour pattern of the algorithm in both RS problems (temperature and wind fields). We have found that once a certain number of stations or nodes goes over a threshold, the improvement in the field reconstruction is obtained by increasing the density of data in the given zones and not by reshaping the number of regions with distant time series. In other words, the algorithm tends to use consistently points of the same areas as N increases. Thus, the performance improvement is obtained by increasing the sampling density in the key regions, rather than including new areas of information. Of course, the obtained zones vary with the variable under study (temperature or wind), but the way in which the algorithm obtains the field reconstruction is similar in both cases. These findings show that the algorithm operates with climatic coherence and validates the methodology exposed herein.

6. The CRO-SL algorithm obtains solutions with climatological consistently in the two RS problems. In the temperature field reconstruction, we found that a set of predefined data points intends to optimally maximize the coverage of the European continent so different temperate and continental climates can be represented within

the reconstruction. These local distributions are similar for both datasets considered (ECA and ERA), giving spatially coherent results. In the wind speed field reconstruction the algorithm mainly includes points over the Northern Atlantic ocean, and additional a few points in the Mediterranean, as those with the most information to reconstruct the wind speed field. This result is clearly explained by the prevailing Westerlies in Europe, which make that points over the Atlantic ocean contain the most information to reconstruct the average monthly wind over the whole continent. Note that, in this case, the best solution obtained does not include any in-land point. Instead, for every solution, all the points are located over the ocean. This is again consistent, since there are no local influences introduced by orography and topography.

7.2 Future research lines

Starting from the significant results obtained from this Ph.D. Thesis, there are several directions in which subsequent studies could progress. Some of the detected areas to be addressed in depth in the near future are the following:

- Due to the generality of the approaches used in the problems tackled in this work, the methodology in this Thesis can be extended to different problems with other atmospheric or oceanic variables. A good example is the study of wave and ocean energy, in which problems of feature selection in prediction problems and significant wave height field reconstruction appear.

- It would be interesting to explore alternative ultra-fast training networks such as CNNs to be hybridized with the CRO-SL. A CNN consists of a number of convolutional and sub-sampling layers optionally followed by fully connected layers. The input of a CNN is usually an image, since their major application up until now is image processing problems. However, there are recent studies which have applied this class of networks to normal (usually very large) databases of classification and regression problems [Gu2018]. The hybridization of CNNs with different versions of the CRO algorithm is therefore a research challenge for the future.

- The application of Machine Learning techniques to palaeo-climatic problems is also an appealing future research line. In fact, the RS tackled in this Thesis can be adapted to represent proxies measurements, so it is possible to obtain the set of proxies which provide the most information for the reconstruction of a given palaeo-climatic variable.

- An innovative research line to be explored is the inclusion of Complex Networks methodologies [Barabasi1999, Albert2002] in problems related to atmospheric sciences. In fact, there are very recent works dealing with somehow related problems in the field of Hydrology [Jha2017b, Naufan2018]. The idea is to obtain evolutionary versions of the networks to

adjust them to a given required pattern, for example to study the relationship of different wind farms production over the time with different characteristics, etc. This idea could also serve to carry out studies related to climate change, by monitoring the changes occurred in the complex networks which model a given phenomenon. This idea has been previously applied to the study of the ENSO and related phenomena [Ludescher2014, Wang2016].

- The last research line to highlight is to apply Soft-Computing related techniques to address Climate Change detection and attribution problems in atmospheric physics. The idea is to model some of this problems as classification, regression or optimization tasks, which can be solved by means of advanced Soft-Computing based approaches. The comparison of these modern computation techniques with classical methodologies and the projection to the future of the obtained results would be alternative lines which could be explored in the future.

Apendix A. List of publications

This Ph.D. Thesis is backed by a number of research articles published in top International Journals:

Papers published in International Journals

1. **S. Salcedo-Sanz**, A. Pastor-Sánchez, A. Blanco-Aguilera, L. Prieto and R. García-Herrera, Feature selection in wind speed prediction systems based on a hybrid Coral Reefs Optimization – Extreme Learning Machine approach. Energy Conversion and Management, 87:10-18, 2014. (JCR: 4.380, Q1).

2. **S. Salcedo-Sanz**, S. Jiménez-Fernández, A. Aybar-Ruíz, C. Casanova-Mateo, J. Sanz-Justo and R. García-Herrera, A CRO-Species optimization scheme for robust global solar radiation statistical downscaling. Renewable Energy, 111:63-76, 2017. (JCR: 4.900, Q1)

3. **S. Salcedo-Sanz**, L. Cornejo-Bueno, L. Prieto, D. Paredes, R. García-Herrera, Feature selection in machine learning prediction systems for renewable energy applications. Renewable and Sustainable Energy Reviews, 90:728-741, 2018. (JCR: 9.184, Q1)

4. **S. Salcedo-Sanz**, R. García-Herrera, C. Camacho-Gómez, A. Aybar-Ruíz and E. Alexandre, Wind speed field reconstruction from a reduced set of representative measuring points. Applied Energy, 228:1111-1121, 2018. (JCR: 7.900, Q1)

Papers Submitted to International Journals

1. **S. Salcedo-Sanz**, R. García-Herrera, C. Camacho-Gómez, E. Alexandre, L. Carro-Calvo and F. Jaume-Santero, Near-optimal selection of representative measuring points for robust temperature field reconstruction with the CRO-SL and Analogue Methods. Global and Planetary Change, Submitted, 2018. (JCR: 3.982, Q1)

Bibliography

[Ahila2015] R. Ahila, V. Sadasivam, K. Manimala, An integrated PSO for parameter determination and feature selection of ELM and its application in classification of power system disturbances. Applied Soft Computing 32:33-37, 2015.

[Ahmed2018] A. S. Ahmed, Wind energy characteristics and wind park installation in Shark El-Ouinat, Egypt. Renewable and Sustainable Energy Reviews 82:734-742, 2018.

[Albert2002] R. Albert, A. L. Barabási, Statistical mechanics of complex networks. Reviews of Modern Physics 74:47-97, 2002.

[Alessandrini2015] S. Alessandrini, L. Delle Monache, S. Sperati, G. Cervone, An analog ensemble for short-term probabilistic solar power forecast. Applied Energy 157:95-110, 2015.

[Alessandrini2015b] S. Alessandrini, L. Delle Monache, S. Sperati, J.N. Nissen, A novel application of an analog ensemble for short-term wind power forecasting. Reneweable Energy 76:768-781, 2015.

[Amorim2012] A. M. Amorim, A. B. Gonçalves, L. Miguel Nunes, A. J. Sousa, Optimizing the location of weather monitoring stations using estimation uncertainty. International Journal of Climatology 32: 941-952, 2012.

[Antonanzas2015] J. Antonanzas, R. Urraca, F. J. Martinez-de-Pison, F. Antonanzas-Torres, Solar irradiation mapping with exogenous data from support vector regression machines estimations. Energy Conversion and Management 100:380-390, 2015.

[Aquila2018] G. Aquila, R. Santana Peruchi, P. Rotela Junior, L. C. Souza Rocha, A R. de Queiroz, E. de Oliveira Pamplona, P. P. Balestrassi, Analysis of the wind average speed in different Brazilian states using the nested GR & R measurement system. Measurement 115:217-222, 2018.

[Aybar2016] A. Aybar-Ruiz, S. Jiménez-Fernández, L. Cornejo-Bueno, C. Casanova-Mateo, J. Sanz-Justo, P. Salvador-González, S. Salcedo-Sanz. A novel Grouping Genetic Algorithm-Extreme Learning Machine approach for global solar radiation prediction from numerical weather models inputs. Solar Energy 132:129-142, 2016.

[Barabasi1999] A. L. Barabási, R. Albert, Emergence of scaling in random networks. Science 286(5439):509-512, 1999.

[Barbounis2007] T.G. Barbounis, J.B. Theocharis, Locally recurrent neural networks for wind speed prediction using spatial correlation. Information Sciences 177:5775-5797, 2007.

[Bauer2015] N. Bauer, V. Bosetti, M. Hamdi-Cherif, A. Kitous, D. McCollum, A. Méjean, S. Rao, H. Turton, L. Paroussos, S. Ashina, et al., CO_2 emission mitigation and fossil fuel markets: Dynamic and international aspects of climate policies. Technological Forecasting and Social Change 90:243-256, 2015.

[Behrang2010] M. A. Behrang, E. Assareh, A. Ghanbarzadeh, A.R. Noghrehabadi, The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Solar Energy 84:1468-1480, 2010.

[Belaid2016] S. Belaid, A. Mellit, Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate. Energy Conversion and Management 118:105-118, 2016.

[Bello2016] G. Bello-Orgaz, J. J. Jung, D. Camacho, Social big data: Recent achievements and new challenges. Information Fusion 28:45-59, 2016.

[Belu2013] R. Belu, Artificial Intelligence Techniques for Solar Energy and Photovoltaic Applications, Handbook of Research on Solar Energy Systems and Technologies, 2013.

[Bermejo2017] E. Bermejo, M. Chica, S. Salcedo-Sanz, O Cordón, Coral Reef Optimization for intensity-based medical image registration. In Proc. of the IEEE Conference on Evolutionary Computation (CEC), 533-540, 2017.

[Bermejo2018] E. Bermejo, M. Chica, S. Damas, S. Salcedo-Sanz, O. Cordón, Coral Reef Optimization with substrate layers for medical Image Registration. Swarm and Evolutionary Computation 42:138-159, 2018.

[Beyer2011] H.G. Beyer, J. Polo Martinez, M. Suri et al., Deliverable 1.1.3. Report on Benchmarking of Radiation Products, Report under contract no. 038665 of MESoR, http://www.mesor.net/deliverables.html, 2011.

[Bhardwaj2013] S. Bhardwaj, V. Sharma, S. Srivastava, O. S. Sastry, B. Bandyopadhyay, S. S. Chandel, J. R. Gupta, Estimation of solar radiation using a combination of Hidden Markov Model and generalized Fuzzy model. Solar Energy 93:43-54, 2013.

[Bianchi2017] E. Bianchi, A. Solarte, T. M. Guozden, Large scale climate drivers for wind resource in Southern South America. Renewable Energy 114:708-715, 2017.

[Bilgili2007] M. Bilgili, B. Sahin, A. Yasar, Application of artificial neural networks for the wind speed prediction of target stations using reference stations data. Renewable Energy 32:2350-2360, 2007.

[Bilgili2011] M. Bilgili, M. Ozoren, Daily total global solar radiation modeling from several meteorological data. Meteorology and Atmospheric Physics 112:125-138, 2011.

[Bisaso2017] K. R. Bisaso, G. T. Anguzu, S. A. Karungi, A. Kiragga, B. Castelnuovo, A survey of machine learning applications in HIV clinical research and care. Computers in Biology and Medicine 91:366-371, 2017.

[Bishop1995] C.M. Bishop, Neural networks for pattern recognition, Oxford University Press, 1995.

[Blum1997] A. Blum, P. Langley, Selection of relevant features and examples in Machine Learning. Artificial Intelligence 97:245-271, 1997.

[Bou2017] M. Bou-Rabee, S. A. Sulaiman, M. S. Saleh, S. Marafi, Using artificial neural networks to estimate solar radiation in Kuwait. Renewable and Sustainable Energy Reviews 72:434-438, 2017.

[Brown2009] S. H. Brown, Multiple linear regression analysis: a matrix approach with MATLAB. Alabama Journal of Mathematics, Spring/Fall, 2009.

[Camacho2018] C. Camacho-Gómez, X. Wang, E. Pereira, I. M. Díaz, S. Salcedo-Sanz, Active vibration control design using the Coral Reefs Optimization with Substrate Layer algorithm. Computers and Structures 157:14-26, 2018.

[Cannon2015] D. J. Cannon, D.J. Brayshaw, J. Methven, P.J. Coker, D. Lenaghan, Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. Renewable Energy 75:767-778, 2015.

[Capellaro2016] M. Capellaro, Prediction of site specific wind energy value factors. Renewable Energy 87(1): 430-436, 2016.

[Carta2015] J. A. Carta, P. Cabrera, J. M. Matías, F. Castellano, Comparison of feature selection methods using ANNs in MCP-wind speed methods. A case study. Applied Energy 158:490-507, 2015.

[Carvalho2014] D. Carvalho, A. Rocha, M. Gómez-Gesteira, C. Silva Santos, Sensitivity of the WRF model wind simulation and wind energy production estimates to planetary boundary layer parameterizations for onshore and offshore areas in the Iberian Peninsula. Applied Energy 135:234-246, 2014.

[Ceylan2004] H. Ceylan, H. K. Ozturk, Estimating energy demand of Turkey based on economic indicators using genetic algorithm approach. Energy Conversion and Management 45:2525-2537, 2004.

[Chardon2014] J. Chardon, B. Hingray, A. C. Favre, P. Autin, J. Gailhard, I. Zin, C. Obled, Spatial similarity and transferability of analog dates for precipitation downscaling over France. Journal of Climate 27:5056-5074, 2014.

[Chen2011] J. L. Chen, H. B. Liu, W. Wu, D. T. Xie, Estimation of monthly solar radiation from measured temperatures using support vector machines - A case study. Renewable Energy 36:413-420, 2011.

[Chidean2015] M. I. Chidean, J. Muõz-Bulnes, J. Ramiro-Bargueño, A. Caamaõ-Fernández, S. Salcedo-Sanz, Spatio-temporal trend analysis of air temperature in Europe and western Asia using data-coupled clustering. Global and Planetary Change 129:45-55, 2015.

[Chimani2013] B. Chimani, C. Matulla, R. Bohm, M. Hofstatter, A new high resolution absolute temperature grid for the Greater Alpine Region back to 1780. International Journal of Climatology 33(9):2129-2141, 2013.

[Coggins2014] J. H. J. Coggins, A. J. McDonald, B. Jolly, Synoptic climatology of the Ross Ice Shelf and Ross Sea region of Antarctica: k-means clustering and validation. International Journal of Climatology 34: 2330-2348, 2014.

[Costa2008] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsend, E. Feitosa, A review on the young history of the wind power short-term prediction. Renewable and Sustainable Energy Reviews 12(6):1725-1744, 2008.

[Dee2011] D. P. Dee, S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, et al., The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society 137:553-597, 2011.

[Deo2017] R. Deo, M. Sahin, Forecasting long-term global solar radiation with an ANN algorithm coupled with satellite-derived (MODIS) land surface temperature (LST) for regional locations in Queensland. Renewable and Sustainable Energy Reviews 72:828-848, 2017.

[Dias2018] S. Díaz, J. A. Carta, J. M. Matías, Performance assessment of five MCP models proposed for the estimation of long-term wind turbine power outputs at a target site using three machine learning techniques. Applied Energy 209:455-477, 2018.

[DMonache2011] L. Delle Monache, T. Nipen, Y. Liu, G. Roux, R. Stull, Kalman filter and analog schemes to postprocess numerical weather predictions. Monthly Weather Review 139:3554-3570, 2011.

[DMonache2013] L. Delle Monache, F. A. Eckel, D. Rife, B. Nagarajan, K. Searight,Probabilistic weather prediction with an Analog ensemble. Monthly Weather Review 141:3498-3516, 2013.

[Deniz2016] N. Deniz, F. Ozcelik, Coral Reefs Optimization algorithm's suitability for dynamic cell formation problems, In Proc. of the Global Joint Conference on Industrial Engineering and Its Application Areas Istanbul, Turkey, June, 2016.

[Diagne2013] M. Diagne, M. David, P. Lauret, J. Boland, N. Schmutz, Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. Renewable and Sustainable Energy Reviews 27:65-76, 2013.

[Diagne2014] M. Diagne, M. David, J. Boland, Post-processing of solar irradiance forecasts from WRF model at Reunion Island. Solar Energy 105:99-108, 2014.

[Dong2014] H. Dong, L. Yang, S. Zhang, Y. Li, Improved prediction approach on solar irradiance of photovoltaic Power Station, TELKOMNIKA Indonesian Journal of Electrical Engineering 12(3)1720-1726, 2014.

[Dorvlo2002] A. S. Dorvlo, J. A. Jervase, A. Al-Lawati, Solar radiation estimation using artificial neural networks. Applied Energy 71:307-319, 2002.

[Duran2017] A. M. Durán-Rosal, D. Guijo-Rubio, P. A. Gutiérrez, S. Salcedo-Sanz, C. Hervás-Martínez, A Coral Reef Optimization algorithm for wave height time series segmentation problems. International Work-Conference on Artificial Neural Networks, 673-684, 2017.

[Duran2018] A. M. Durán-Rosal, P. A. Gutiérrez, S. Salcedo-Sanz, C. Hervás-Martínez, A statistically-driven Coral Reef Optimization algorithm for optimal size reduction of time series. Applied Soft Computing 63:139-153, 2018.

[Eiben2003] A. E. Eiben, J. E. Smith, Introduction to evolutionary computing, Springer-Verlag, Berlin, 2003.

[Eiben2015] A. E. Eiben, J. Smith, From evolutionary computation to evolution of things. Nature 52(5):476-482, 2015.

[ELM2018] Huang GB. ELM matlab code. http://www.ntu.edu.sg/home/egbhuang/elm_codes.html; [accessed 15 February 2018].

[Falkenauer1992] E. Falkenauer, The grouping genetic algorithm–Widening the scope of the GAs. Belgian Journal of Operations Research, Statistics and Computer Science 33:79-102, 1992.

[Feng2017] C. Feng, M. Cui, B. M. Hodge, J. Zhang, A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Applied Energy 190:1245-1257, 2017.

[Ferreira2014] A. J. Ferreira, M. T. Figueiredo, Incremental filter and wrapper approaches for feature discretization. Neurocomputing 123:60-74, 2014.

[Ficco2016] M. Ficco, C. Esposito, F. Palmieri, A. Castiglione, A Coral-Reefs and game theory-based approach for optimizing elastic cloud resource allocation, Future Generation Computer System 78:343-352, 2018.

[François2017] B. François, S. Martino, L. S. Tøfte, B. Hingray, B. Mo, J. D. Creutin, Effects of increased wind power generation on mid-Norway's energy balance under climate change: a market based approach. Energies 10(227):1-18, 2017.

[Fu2013] C. L. Fu, H. Y. Cheng, Predicting solar irradiance with all-sky image features via regression. Solar Energy 97:537-550, 2013.

[Garcia2018] O. García-Hinde, G. Terrén-Serrano, M. A. Hombrados-Herrera, V. Gómez-Verdejo, S. Jiménez-Fernández, C. Casanova-Mateo et al., Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting. Engineering Applications of Artificial Intelligence 69:157-167, 2018.

[Gardner1998] M.W Gardner, S.R Dorling, Artificial Neural Networks (the multilayer perceptron)- A review of applications in the atmospheric sciences. Atmospheric Environment 32(14/15):2627-2636, 1998.

[Geem2001] Z. W. Geem, J. H. Kim, G. V. Loganathan, A new heuristic optimization algorithm: Harmony Search. Simulation 76(2):60-68, 2001.

[Gibergans2007] J. Gibergans-Báguena, M. C. Llasat, Improvement of the analog forecasting method by using local thermodynamic data. Application to autumn precipitation in Catalonia. Atmospheric Research 86:173-193, 2007.

[Gu2018] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, et al., Recent advances in convolutional neural networks. Pattern Recognition 77:354-377, 2018.

[Gupta2011] R. A. Gupta, R. Kumar, A. K. Bansal, Selection of input variables for the prediction of wind speed in wind farms based on genetic algorithms. Wind Energy 35(6):649-660, 2011.

[Hagan1994] M.T. Hagan, M.B. Menhaj, Training feed forward network with the Marquardt algorithm. IEEE Transactions on Neural Networks 5(6):989-993, 1994.

[Haykin1998] S. Haykin, Neural networks: a comprehensive foundation, Cambridge University Press, Cambridge, UK, 1998.

[Heinermann2016] J. Heinermann, O. Kramer, Machine learning ensembles for wind power prediction. Renewable Energy 89:671-679, 2016.

[Hocaoglu2008] F. O. Hocaoglu, O. N. Gerek, M. Kurban, Hourly solar radiation forecasting using optimal coefficient 2-D linear filters and feed-forward neural networks. Solar Energy 82:714-726, 2008.

[Hocaoglu2009] F. O. Hocaoglu, M. Fidan, O. N. Gerek, Mycielski approach for wind speed prediction. Energy Conversion and Management 50(6):1436-1443, 2009.

[Horton2017] P. Horton, M. Jaboyedoff, C. Obled, Global optimization of an Analog Method by means of Genetic Algorithms. Monthly Weather Review 145:1275-1294, 2017.

[Hu2015] Z. Hu, Y. Bao, R. Chiong, T. Xiong, Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection. Energy 84:419-431, 2015.

[Huang2006] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications. Neurocomputing 70(1):489-501, 2006.

[Huang2012] G.B. Huang, H. Zhou, X. Ding, R. Zhang, Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 42(2):513-519, 2012.

[Huang2015] G. Huang, G.B. Huang, S. Song, K. You, Trends in extreme learning machines: A review. Neural Networks 61:32-48, 2015.

[Huda2014] S. Huda, M. Abdollahian, M. Mammadov, J. Yearwood, S. Ahmed, I. Sultan, A hybrid wrapper-filter approach to detect the source(s) of out-of-control signals in multivariate manufacturing process. European Journal of Operatinal Research 237:857-870, 2014.

[Huda2016] S. Huda, J. Abawajy, M. Alazab, M. Abdollalihian, R. Islam, J. Yearwood, Hybrids of support vector machine wrapper and filter based framework for malware detection. Future Generation Computer Systems 55:376-390, 2016.

[Inman2013] R. H. Inman, H. T. Pedro, C. F. Coimbra, Solar forecasting methods for renewable energy integration. Progress in Energy and Combustion Science 39(6):535-576, 2013.

[Jha2017] S. K. Jha, J. Bilalovic, A. Jha, N. Patel, H. Zhang, Renewable energy: Present research and future scope of Artificial Intelligence. Renewable and Sustainable Energy Reviews 77:297-317, 2017.

[Jha2017b] S. K. Jha, B. Sivakumar, Complex networks for rainfall modeling: Spatial connections, temporal scale, and network size. Journal of Hydrology 554:482-489, 2017.

[Jiang2017] Y. Jiang, G. Huang, Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction. Energy Conversion and Management 2017;144:340-350, 2017.

[Jiang2011] Y. Jiang, B. Tang, Y. Qin, W. Liu, Feature extraction method of wind turbine based on adaptive Morlet wavelet and SVD. Renewable Energy 36:2146-2153, 2011.

[Jiang2013] Y. Jiang, Z. Song, A. Kusiak, Very short-term wind speed forecasting with Bayesian structural break model. Renewable Energy 50:637-647, 2013.

[Jinming2015] F. Jinming, L. Yonghe, Y. Zhongwei. Analysis of surface air temperature warming rate of China in the last 50 years (1962-2011) using k-means clustering. Theoretical and Applied Climatology 120:785-796, 2015.

[John1994] G. John, R. Kohavi, K. Pfleger, Irrelevant features and the subset selection problem. Proceedings of the 11th International Conference on Machine Learning; 1994 July 10-13; New Brunswick, NJ, USA, 1994.

[Jones1998] D. R. Jones, M. Schonlau, W. J. Welch. Efficient global optimization of expensive black-box functions. Journal of Global Optimization 13(4): 455-492, 1998.

[Jones2017] L.E. Jones, Practical Management of Variability, Uncertainty, and Flexibility in Power Grid. Renewable Energy Integration, Academic Press: Cambridge, MA, USA, 2017.

[Junk2015] C. Junk, L. Delle Monach, S. Alessandrini, Analog-based ensemble model output statistics. Monthly Weather Review 143:2909-2916, 2015.

[Jurado15] S. Jurado, A. Nebot, F. Múgica, N. Avellana, Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques. Energy 86:276-291, 2015.

[Jursa2007] R. Jursa, Variable selection for wind power prediction using particle swarm optimization. In Proceedings of the 9th Genetic and Evolutionary Computation Conference (GECCO), London, England, 2059-2065, 2007.

[Jursa2008] R. Jursa, K. Rohrig, Short-term wind power forecasting using evolutionary algorithms for the automated specification of artificial intelligence models. International Journal of Forecasting 24:694-709, 2008.

[Kalogirou2006] S. Kalogirou, Artificial Intelligence in Energy and Renewable Energy Systems. Nova Publishers, NY, USA, 2006.

[Kalogirou2014] S. A. Kalogirou, Designing and Modeling Solar Energy Systems, Solar Energy Engineering, (Second Edition), Chapter 11, 583-699, 2014.

[Kanamitsu1991] M. Kanamitsu, J.C. Alpert, K.A. Campana, P.M. Caplan, D.G. Deaven, M. Iredell, B. Katz, H.L. Pan, J. Sela, G.H. White, Recent changes implemented into the Global Forecast System at NMC. Weather and Forecasting 6(3):425-436, 1991.

[Katrin2012] S. Katrin, C. Philippe, B. Reinald, M. Bert, H. Cora, R. Jacqueline, Automatic selection of a representative trial from multiple measurements using Principle Component Analysis. Journal of Biomechanics 45:2306-2309, 2012.

[Kennedy1995] J. Kennedy, R. Eberhart. Particle swarm optimization. Proc. of the 4th IEEE International Conference on Neural Networks 1942-1948, 1995.

[Khatib2012] T. Khatib, A. Mohamed, K. Sopian, A review of solar energy modeling techniques. Renewable and Sustainable Energy Reviews 16:2864-2869, 2012.

[Khashei2009] M. Khashei, M. Bijari, G. Raissi-Ardali, Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs). Neurocomputing 72:956-967, 2009.

[Kiran2012] M. S. Kiran, E. Özceylan, M. Gündüz, T. Paksoy, A novel hybrid approach based on Particle Swarm Optimization and Ant Colony Optimization to forecast energy demand of Turkey. Energy Conversion and Management 53:75-83, 2012.

[Kirchner2013] N. Kirchner-Bossi, L. Prieto, R. García-Herrera, L. Carro-Calvo, S. Salcedo-Sanz, Multi-decadal variability in a centennial reconstruction of daily wind. Applied energy 105:30-46, 2013.

[Kirchner2015] N. Kirchner-Bossi, R. García-Herrera, L. Prieto, R. M. Trigo, A long-term perspective of wind power output variability. International Journal of Climatology 35 (9):2635-2646, 2015.

[Kirpatrick1983] D. Kirpatrick, C. D. Gerlatt, M. P. Vecchi, Optimization by simulated annealing. Science 220:671-680, 1983.

[Klein2002] A. M. Klein-Tank, et al., Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment. International Journal of Climatology 22(12):1441-1453, 2002.

[Kleissl2013] J. Kleissl (ed), Solar energy forecasting and resource assessment, Academic Press, 2013.

[Kohavi1997] R. Kohavi, G. John, Wrappers for features subset selection. International Journal of Digital Libraries 1:108-121, 1997.

[Kong2015] X. Kong, X. Liu, R. Shi, K. Y. Lee, Wind speed prediction using reduced support vector machines with feature selection. Neurocomputing 169:449-456, 2015.

[Koprinska2015] I. Koprinska, M. Rana, V. G, Agelidis, Correlation and instance based feature selection for electricity load forecasting. Knowledge-Based Systems 82:29-40, 2015.

[Kottek2006] M. Kottek, J. Grieser, C. Beck, B. Rudolf, F. Rubel, World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 15:259-263, 2006.

[Kou2014] P. Kou, D. Liang, F. Gao, L. Gao, Probabilistic wind power forecasting with online model selection and warped Gaussian Process. Energy Conversion and Management 84:649-663, 2014.

[Kumar2015] S. Kumar, D. López, Feature selection used for wind speed forecasting with data driven approaches. Journal of Engineering Science and Technological Review 8:124-127, 2015.

[Kumar2016] Y. Kumar, J. Ringenberg, S. S. Depuru, V. K. Devabhaktuni, J. W. Lee, E. Nikolaidis, B. Andersen, A. Afjeh, Wind energy: Trends and enabling technologies, Renewable and Sustainable Energy Reviews 53:209-224, 2016.

[Kusiak2010b] A. Kusiak, W. Li, Short-term prediction of wind power with a clustering approach. Renewable Energy 35(10):2362-2369, 2010.

[Kusiak2010] A. Kusiak, W. Li, Estimation of wind speed: A data-driven approach. Journal of Wind Engineering and Industrial Aerodynamics 98:559-567, 2010.

[Landberg1999] L. Landberg, Short-term prediction of the power production from wind farms. Journal of Wind Engineering and Industrial Aerodynamics 80:207-220, 1999.

[Landberg2001] L. Landberg, Short-term prediction of local wind conditions. Journal of Wind Engineering and Industrial Aerodynamics 89:235-245, 2001.

[LeCun2015] Y. LeCun, B. Yoshua, G. Hinton, Deep learning. Nature 521.7553:436-444, 2015.

[Lguensat2017] R. Lguensat, P. Tandeo, P. Ailliot, M. Pulido, R. Fablet, The analog data assimilation. Monthly Weather Review 145(10):4093-4107, 2017.

[Li2001] S. Li, C. D. Wunch, A. O'Hair, G. M. Giesselmann, Using neural networks to estimate wind turbine power generation. IEEE Transactions on Energy Conversion 16(3):276-281, 2001.

[Li2010] G. Li, J. Shi, On comparing three artificial neural networks for wind speed forecasting. Applied Energy 87(7):2313-2320, 2010.

[Li2016] M. Li, C. Miao, C. Leung, A Coral Reef Algorithm based on learning automata for the coverage control problem of heterogeneous directional sensor networks. Sensors 15:30617-30635, 2015.

[Lima2016] F. J. Lima, F. R. Martins, E. B. Pereira, E. Lorenz, D. Heinemann, Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks, Renewable Energy 87:807-818, 2016.

[Lin1996] T. Lin, B. G. Horne, P. Tino, C. L. Giles, Learning long-term dependencies in NARX recurrent neural networks. IEEE Transactions on Neural Networks 7:1329-1351, 1996.

[Litta2012] A.J. Litta, U.C. Mohanty, S.M. Idicula, The diagnosis of severe thunderstorms with high-resolution WRF model, Journal of Earth System Science 121(2):297-316, 2012.

[Liu2012] H. Liu, H. Tian, Y. F. Li, Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. Applied Energy 98:415-424, 2012.

[Liu2013] H. Liu, H. Tian, D. Pan, Y. Li, Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. Applied Energy 107:191-208, 2013.

[Lopez2005] G. López, B. Batlles, J. Tovar-Pescador, Selection of input parameters to model direct solar irradiance by using artificial neural networks. Energy 30:1675-1684, 2005.

[Lorentz1969] E. Lorentz, Atmospheric predictability as revealed by naturally occurring analogues. Journal of Atmospheric Science 26:636-646, 1969.

[Lozier2008] M. S. Lozier, The Spatial Pattern and Mechanisms of Heat-Content Change in the North Atlantic. Science 319:800, 2008.

[Ludescher2014] J. Ludescher, A. Gozolchiani, M. I. Bogachev, A. Bunde, S. Havlin, H. J. Schellnhuber. Very early warning of next El Niño. Proceedings of the national Academy of sciences 111(6):2064-2066, 2014.

[Lutz2016] F. Lutz, H. W. ter Maat, H. Biemans, A. B. Shrestha, P. Westerdan, W. W. Immerzeel, Selecting representative climate models for climate change impact studies: an advanced envelope-based selection approach. International Journal of Climatology 36:3988-4005, 2016.

[Medeiros2015] I. G. Medeiros, J. C. Xavier-Júnior, A. M. Canuto, Applying the Coral Reefs Optimization algorithm to clustering problems, In Proc. of the International Joint Conference on Neural Networks (IJCNN), 1-8, 2015.

[Mellit2008] A. Mellit, S. A. Kalogirou, Artificial intelligence techniques for photovoltaic applications: a review. Progress in Energy and Combustion Science 34:574-632, 2008.

[Michalewicz2000] Z. Michalewicz, D. B. Fogel, How to solve it: modern heuristics, Springer-Verlag, Berlin, 2000.

[Mohandes2004] M. A. Mohandes, T. O. Halawani, S. Rehman, A. A. Hussain, Support vector machines for wind speed prediction. Renewable Energy 29:939-947, 2004.

[Mohammadi2015] K. Mohammadi, S. Shamshirband, C. W. Tong, M. Arif, D. Petkovic, S. Che, A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation. Energy Conversion and Management 92:162-171, 2015.

[Mohammadi2016] K. Mohammadi, S. Shamshirband, D. Petkovic, H. Khorasanizadeh, Determining the most important variables for diffuse solar radiation prediction using adaptive neuro-fuzzy methodology; case study: City of Kerman, Iran. Renewable and Sustainable Energy Reviews 53:1570-1579, 2016.

[Monteiro2017] R. V. Monteiro, G. C. Guimarães, F. A. Moura, M. R. Albertini, M. K. Albertini, Estimating photovoltaic power generation: Performance analysis of artificial neural networks, Support Vector Machine and Kalman filter. Electric Power Systems Research 143:643-656, 2017.

[Munteanu2016] I. Munteanu, G. Besancon, Identification-based prediction of wind park power generation. Renewable Energy 97: 422-433, 2016.

[Nasseri2011] M. Nasseri, B. Zahraie, Application of simple clustering on space-time mapping of mean monthly rainfall pattern. International Journal of Climatology 31: 732-741, 2011.

[Naufan2018] I. Naufan, B. Sivakumar, F. M. Woldemeskel, S. V. Raghavan, M. T. Vu, S. Y. Liong, Spatial connections in regional climate model rainfall outputs at different temporal scales: Application of network theory. Journal of Hydrology 556:1232-1243, 2018.

[Norgaard2004] P. Norgaard, H. Holttinen, A multi-turbine power curve approach, In Proc. of the Nordic Wind Power Conference 1-2, 1-5, 2004.

[Olatomiwa2015] L. Olatomiwa, S. Mekhilef, S. Shamshirband, K. Mohammadi, D. Petkovic, S. Ch, A support vector machine-firefly algorithm-based model for global solar radiation prediction, Solar Energy 115:632-644, 2015.

[Ortiz2011] E. G. Ortiz-García, S. Salcedo-Sanz, A. M. Pérez-Bellido, J. Gascón-Moreno, J. A. Portilla-Figueras, L. Prieto, Short-term wind speed prediction in wind farms based on banks of support vector machines. Wind Energy 14(2):193-207, 2011.

[PerezO2016] M. Pérez-Ortiz, S. Jiménez-Fernández, P. A. Gutiérrez, E. Alexandre, C. Hervás-Martínez, S. Salcedo-Sanz, A review of classification problems and algorithms in renewable energy applications. Energies 9:1-27, 2016.

[Peters2013] G.P. Peters, R.M. Andrew, T. Boden, J.G. Canadell, P. Ciais, C. Le Quéré, G. Marland, M.R. Raupach, C. Wilson, The challenge to keep global warming below 2°C. Nature Climate Change 3:4-6, 2013.

[Pichpibul12] T. Pichpibul, R. Kawtummachai A modified Coral-Reef Optimization algorithm for the capacitated vehicle routing problem, In Proceedings of the 29th International Technical Conference on Circuit/Systems Computers and Communications (ITC-CSCC) Phuket, Thailand, 684-687, 2014.

[Pichpibul12b] T. Pichpibul, R. Kawtummachai, An improved Clarke and Wright savings algorithm for the capacitated vehicle routing problem. Science Asia 38:307-318, 2012.

[Poorter2017] E. De Poorter, T. Van Haute, Eric Laermans, Ingrid Moerman, Benchmarking of localization solutions: guidelines for the selection of evaluation points. Ad Hoc Networks 59:86-96, 2017.

[Pourmousavi2011] S.A. Pourmousavi Kani, M.M. Ardehali, Very short-term wind speed prediction: A new artificial neural network-Markov chain model. Energy Conversion and Management 52(1):738-745, 2011.

[Pryor2005] S. C. Pryor, J. T. Schoof, R. J. Barthelmie, Climate change impacts on wind speeds and wind energy density in northern Europe: empirical downscaling of multiple AOGCMs. Climate Research 29:183-198, 2005.

[Pryor2006] S.C. Pryor, R.J. Barthelmie, J. T. Schoof, Inter-annual variability of wind indices across Europe. Wind Energy 9:27-38, 2006.

[Pryor2010] S. C. Pryor, R. J. Barthelmie, Climate change impacts on wind energy: A review. Renewable and Sustainable Energy Reviews 14:430-437, 2010.

[Pryor2011] S. C. Pryor, R. J. Barthelmie, Assessing climate change impacts on the near-term stability of the wind energy resource over the United States, Proceedings of the National Academy of Sciences 108(20):8167-8171, 2011.

[Qiao2013] J. Zeng, W. Qiao, Short-term solar power prediction using a support vector machine. Renewable Energy 52:118-127, 2013.

[Rana2016] M. Rana, I. Koprinska, V. G. Agelidis, Univariate and multivariate methods for very short-term solar photovoltaic power forecasting. Energy Conversion and Management 121:380-390, 2016.

[Rehman2008] S. Rehman, M. Mohandes, Artificial neural network estimation of global solar radiation using air temperature and relative humidity. Energy Policy 36(2):571-576, 2008.

[REN21-2017] REN21, Renewables 2017 Global Status Report, Paris, 2017.

[Renani2016] E. T. Renani, M.F.M. Elias, N.A. Rahim, Using data-driven approach for wind power prediction: A comparative study. Energy Conversion and Management 118:193-203, 2016.

[Riahy2008] G.H. Riahy, M. Abedi, Short term wind speed forecasting for wind turbine applications using linear prediction. Renewable Energy 33(1):35-41, 2008.

[Rife2013] S. L. Rife, E. Vanvyve, J. O. Pinto, A. J. Monaghan, C. A. Davis, G. S. Poulos. Selecting representative days for more efficient dynamical climate downscaling: application to wind energy. Journal of Applied Meteorology and Climatology 52:47-63, 2013.

[Ringkjob2017] H. K. Ringkjob, B. Jourdier, P. Drobinski, R. Plougonven, P. Tankov, Modelling the variability of the wind energy resource on monthly and seasonal timescales. Renewable Energy 113:1434-1446, 2017.

[Ritter2017] M. Ritter, L. Deckert, Site assessment, turbine selection, and local feed-in tariffs through the wind energy index. Applied Energy 185(2):1087-1099, 2017.

[Robert2013] S. Robert, L. Foresti, M. Kanevski, Spatial prediction of monthly wind speeds in complex terrain with adaptive general regression neural networks. International Journal of Climatology 33:1793-1804, 2013.

[Rowland2016] C. S. Rowland, J. W. Mjelde, Politics and petroleum: Unintended implications of global oil demand reduction policies. Energy Research and Social Science 11:209-224, 2016.

[Ruane2017] A. C. Ruane, S. P. McDermid, Selection of a representative subset of global climate models that captures the profile of regional changes for integrated climate impacts assessment. Earth Perspective 4(1):1-20, 2017.

[Saavedra2013] B. Saavedra-Moreno, S. Salcedo-Sanz, L. Carro-Calvo, J. Gascón-Moreno, S. Jiménez-Fernández, L. Prieto, Very fast training neural-computation techniques for real measure-correlate-predict wind operations in wind farms. Journal of Wind Engineering and Industrial Aerodynamics 116:49-60, 2013.

[Sahin2013] M. Sahin, Y. Kaya, M. Uyar, S. Yidirim, Application of extreme learning machine for estimating solar radiation from satellite data. International Journal of Energy Research 38(2): 205-212, 2014.

[Salcedo2002] Salcedo-Sanz S, Prado-Cumplido M, Pérez-Cruz F, Bousoño-Calzón C. Feature selection via genetic optimization. International Conference on Artificial Neural Networks, International Conference on Artificial Neural Networks, ICANN2002, 2002.

[Salcedo2009] S. Salcedo-Sanz, A.M. Pérez-Bellido, E.G. Ortiz-García, A. Portilla-Figueras, L. Prieto, D. Paredes, Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction. Renewable Energy 34:1451-1457, 2009.

[Salcedo2009b] S. Salcedo-Sanz, A. M. Pérez-Bellido, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, F. Correoso, Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks. Neurocomputing 72:1336-1341, 2009.

[Salcedo2009c] S. Salcedo-Sanz, A survey of repair methods used as constraint handling techniques in evolutionary algorithms. Computer Science Reviews 3:175-192, 2009.

[Salcedo2011] S. Salcedo-Sanz, E. G. Ortiz-García, A. M. Pérez-Bellido, A. Portilla-Figueras, L. Prieto, Short term wind speed prediction based on evolutionary support vector regression algorithms. Expert Systems with Applications 38(4): 4052-4057, 2011.

[Salcedo2013] S. Salcedo-Sanz,C. Casanova-Mateo, A. Pastor-Sánchez, D. Gallo-Marazuela, A. Labajo-Salazar, A. Portilla-Figueras, Direct solar radiation prediction based on Soft-Computing algorithms including novel predictive atmospheric variables, Intelligent Data Engineering and Automated Learning - IDEAL 2013, Lecture Notes in Computer Science, 8206:318-325, 2013.

[Salcedo2014] S. Salcedo-Sanz, D. Gallo-Marazuela, A. Pastor-Sánchez, L. Carro-Calvo, A. Portilla-Figueras, L. Prieto, Offshore wind farm design with the Coral Reefs Optimization algorithm. Renewable Energy 63:109-115, 2014.

[Salcedo2014b] S. Salcedo-Sanz, J.L. Rojo, M. Martínez-Ramón, G. Camps-Valls, Support vector machines in engineering: an overview. WIREs Data Mining and Knowledge Discovery 4(3):234-267, 2014.

[Salcedo2014c] S. Salcedo-Sanz, C. Casanova-Mateo, J. Muñoz-Marí, G. Camps-Valls, Prediction of daily global solar irradiation using temporal Gaussian processes, IEEE Geoscience and Remote Sensing Letters 11(11):1936-1940, 2014.

[Salcedo2014d] S. Salcedo-Sanz, J. Del Ser, I. Landa-Torres, S. Gil-López, and J. A. Portilla-Figueras, The Coral Reefs Optimization algorithm: a novel metaheuristic for efficiently solving optimization problems. The Scientific World Journal 739768:1-15, 2014.

[Salcedo2014e] S. Salcedo-Sanz, C. Casanova-Mateo, A. Pastor-Sánchez, M. Sánchez-Girón, Daily global solar radiation prediction based on a Hybrid Coral Reefs Optimization – Extreme Learning Machine approach. Solar Energy 105:91-98, 2014.

[Salcedo2014f] S. Salcedo-Sanz, J. E. Sánchez-García, J. A. Portilla-Figueras, S. Jiménez-Fernández, A. M. Ahmadzadeh, A Coral-Reefs Optimization algorithm for the optimal service distribution problem in mobile radio access networks. Transactions on Emergent Telecommunication Technologies 25(11):1057-1069, 2014.

[Salcedo2014g] Salcedo-Sanz, P. García-Díaz, J.A. Portilla-Figueras, J. Del Ser, S. Gil-López, A Coral Reefs Optimization algorithm for optimal mobile network deployment with electromagnetic pollution control criterion. Applied Soft Computing 24:239-248, 2014.

[Salcedo2015] S. Salcedo-Sanz, J. Muñoz-Bulnes, J. A. Portilla-Figueras, J. del Ser, One-year-ahead energy demand estimation from macroeconomic variables using Computational Intelligence algorithms. Energy Conversion and Management 99:62-71, 2015.

[Salcedo2015b] S. Salcedo-Sanz, A. Pastor-Sánchez, J. Del Ser, L. Prieto, Z.W. Geem, A Coral Reefs Optimization algorithm with Harmony Search operators for accurate wind speed prediction. Renewable Energy 75:93-101, 2015.

[Salcedo2015c] S. Salcedo-Sanz, A. Pastor-Sánchez, J. Del Ser, L. Prieto, Z. W. Geem. A Coral Reefs Optimization algorithm with Harmony Search operators for accurate wind speed prediction. Renewable Energy 75:93-101, 2015.

[Salcedo2016] S. Salcedo-Sanz, C. Camacho-Gómez, R. Mallol-Poyato, S. Jiménez-Fernández, J. Del Ser, A novel Coral Reefs Optimization algorithm with substrate layers for optimal battery scheduling optimization in micro-grids. Soft Computing 20(11):4287-4300, 2016.

[Salcedo2016b] Salcedo-Sanz, P. García-Díaz, J. Del Ser, M. N. Bilbao, J. A. Portilla-Figueras, A novel Grouping Coral Reefs Optimization algorithm for optimal mobile network deployment problems under electromagnetic pollution and capacity control criteria. Expert Systems with Applications 55:388-2402, 2016.

[Salcedo2016c] S. Salcedo-Sanz, Modern meta-heuristics based on nonlinear physics processes: A review of models and design procedures. Physics Reports 655:1-70, 2016.

[Salcedo2017] S. Salcedo-Sanz, C. Camacho-Gómez, A. Magdaleno, E. Pereira, A. Lorenzana, Structures vibration control via tuned mass dampers using a co-evolution coral reefs optimization algorithm. Journal of Sound and Vibration 393:62-75, 2017.

[Salcedo2017b] S. Salcedo-Sanz, J. Muñoz-Bulnes, M. Vermeij, New Coral Reefs-based Approaches for the Model Type Selection Problem: A Novel Method to Predict a Nation's Future Energy Demand. International Journal of Bio-inspired Computation 10(3):145-158, 2017.

[Schmidli2007] J. Schmidli, C. M. Goodess, C. Frei, M. R. Haylock, Y. Hundecha, J. Ribalaygua, T. Schmith, Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. Journal of Geophysical Research 112:1-20, 2007.

[Senkal2009] O. Senkal, T. Kuleli, Estimation of solar radiation over Turkey using artificial neural network and satellite data. Applied Energy 86(7-8):1222-1228, 2009.

[Serrano2014] J. Serrano-González, M. Burgos-Payán, J. M. Riquelme-Santos, F. González-Longatt, A review and recent developments in the optimal wind-turbine micro-siting problem, Renewable and Sustainable Energy Reviews 30:133-144, 2014.

[Sgubin2017] G. Sgubin et al. Abrupt cooling over the North Atlantic in modern climate models. Nature Communications 8:14375, 2017.

[Sharma2016] V. Sharma, D. Yang, W. Walsh, T. Reindl, Short term solar irradiance forecasting using a mixed wavelet neural network. Renewable Energy 90: 481-492, 2016.

[Sharma2018] A. Sharma, A. Kakkar, Forecasting daily global solar irradiance generation using machine learning. Renewable and Sustainable Energy Reviews 82:2254-2269, 2018.

[Sharp2018] M. Sharp, R. Ak, T. Hedberg, A survey of the advancing use and development of machine learning in smart manufacturing, Journal of Manufacturing Systems 48:170-179, 2018.

[Silva2016] H. M. Silva, A. M. Canuto, Inácio G. Medeiros, J. C. Xavier-Júnior, Cluster ensembles optimization using the Coral reefs Optimization Algorithm, Artificial Neural Networks and Machine Learning – ICANN16, Lecture Notes in Computer Science 9887:275-282, 2016.

[Simmons1989] A.J. Simmons, D.M. Burridge, M. Jarraud, C. Girard, W. Wergen, The ECMWF medium-range prediction models: development of the numerical formulations and the impact of increased resolution. Meteorology and Atmospheric Physics 40:28-60, 1989.

[Skamarock2005] W.C. Skamarock, J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, W. Wang, J.G. Powers, A Description of the Advanced Research WRF Version 2, Technical Report; Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research: Boulder, CO, USA, 2005.

[Skamarock2008] W.C. Skamarock, J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.Y. Huang, W. Wang, W.G. Powers, A description of the advanced research WRF, Version 3, NCAR Technical Note, 2008.

[Smola2004] A. J. Smola, B. Schölkopf, A tutorial on support vector regression. Statistics and Computing 14:199-222, 2004.

[Solorio2016] S. Solorio-Fernández, J. A. Carrasco-Ochoa, J. F. Martínez-Trinidad, A new hybrid filter-wrapper feature selection method for clustering based on ranking. Neurocomputing 214:866-880, 2016.

[Souloukngа2018] M. H. Souloukngа, S.Y. Doka, N.Revanna, N.Djongyang, T. C. Kofane, Analysis of wind speed data and wind energy potential in Faya-Largeau, Chad, using Weibull distribution. Renewable Energy 121:1-8, 2018.

[Sozen2004] A. Sozen, E. Arcakliogblu, M. Ozalp, Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data. Energy Conversion and Management 45:3033-3052, 2004.

[Storn1997] R. Storn, K. Price. Differential Evolution - A simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11:341-359, 1997.

[Suganthi2012] L. Suganthi, A. A. Samuel, Energy models for demand forecasting – A review. Renewable and Sustainable Energy Reviews 16:1223-1240, 2012.

[Tar2008] K. Tar, Some statistical characteristics of monthly average wind speed at various heights. Renewable and Sustainable Energy Reviews 12(6):1712-1724, 2008.

[Tasci2014] A. Tascikaraoglu, M. Uzunoglu, A review of combined approaches for prediction of short-term wind speed and power. Renewable and Sustainable Energy Reviews 34:243-254, 2014.

[Torkkola2000] K. Torkkola, W. M. Campbell, Mutual information in learning feature transformations. Proceedings of the 17th International Conference on Machine Learning, San Francisco, CA., 1015-1022, 2000.

[Torkkola2002] K. Torkkola, On feature extraction by mutual information maximization. Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 821-824, 2002.

[Torres2005] J.L. Torres, A. García, M. De Blas, A. De Francisco. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). Solar Energy 79:65-77, 2005.

[Vanvyve2015] E. Vanvyve, L. Delle Monache, A. J. Monaghan, J. O. Pinto, Wind resource estimates with an analog ensemble approach. Renewable Energy 74: 761-773, 2015.

[Vermeij2005] M. J. Vermeij, Substrate composition and adult distribution determine recruitment patterns in a Caribbean brooding coral. Marine Ecology Progress Series 295:123-133, 2005.

[Voyant2011] C. Voyant, M. Muselli, C. Paoli, M. L. Nivet, Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation. Energy 36(1):348-359, 2011.

[Voyant2013] C. Voyant, M. Muselli, C. Paoli, M. L. Nivet, Hybrid methodology for hourly global radiation forecasting in Mediterranean area, Renewable Energy 53:1-11, 2013.

[Voyant2017] C. Voyant, G. Notton, S. Kalogirou, M. L. Nivet, C. Paoli, F. Motte, A. Fouilloy, Machine learning methods for solar radiation forecasting: A review. Renewable Energy 105:569-582, 2017.

[Wallace2006] J. Wallace, P. Hobbs, Atmospheric Science - An Introductory Survey, Elsevier, 2006.

[Wang2015] F. Wang, Z. Zhen, Z. Mi, H. Sun, S. Su, G. Yang, Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. Energy and Buildings 86:427-438, 2015.

[Wang2016] Y. Wang, A. Gozolchiani, Y. Ashkenazy, S. Havlin, Oceanic El-Niño wave dynamics and climate networks. New Journal of Physics 18(3):033021, 2016.

[Wang2017] H. Wang, Y. Kawahara, C. Weng, J. Yuan. Representative selection with structured sparsity. Pattern Recognition 63:268-278, 2017.

[Weston2000] J. Weston, S. Mukherjee, O. Chapelle, M. Pontil, T. Poggio, V. Vapnik, Feature selection for SVMs. NIPS'00 Proceedings of the 13th International Conference on Neural Information Processing Systems, Denver, CO. MIT Press, Cambridge, MA, USA, 2000.

[Wilcke2016] R. A. Wilcke, L. Bärring, Selecting regional climate scenarios for impact modelling studies. Environmental Modeling and Software 78:191-201, 2016.

[Will2011] A.Will, J. Bustos, M. Bocco, J. Gotaya, C. Lamelas, On the use of niching genetic algorithms for variable selection in solar radiation estimation. Renewable Energy 50:168-176, 2011.

[Wu2016] Y. Wu, J. Wang, A novel hybrid model based on artificial neural networks for solar radiation prediction, Renewable Energy 89:268-284, 2016.

[Wu2014] J. Wu, C. K. Chan, Y. Zhang, B. Y. Xiong, Q. H. Zhang, Prediction of solar radiation with genetic approach combing multi-model framework. Renewable Energy 66:132-139, 2014.

[Wu2011] J. Wu, C. K. Chan, Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN. Solar Energy 85:808-817, 2011.

[Yacef2012] R. Yacef, M. Benghanem, A. Mellit, Prediction of daily global solar irradiation data using Bayesian neural network: A comparative study. Renewable Energy 48:146-154, 2012.

[Yadav2014] A. K. Yadav, S. S. Chandel, Solar radiation prediction using Artificial Neural Network techniques: A review. Renewable and Sustainable Energy Reviews 33:772-781, 2014.

[Yadav2014b] A. K. Yadav, H. Malik, S. S. Chandel, Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models. Renewable and Sustainable Energy Reviews 31:509-519, 2014.

[Yan2015] J. Yan, Y. Liu, S. Han, Y. Wang, S. Feng Reviews on uncertainty analysis of wind power forecasting. Renewable and Sustainable Energy Reviews 52:1322-1330, 2015.

[Yang1998] Yang J, Honavar V. Feature subset selection using a genetic algorithm. IEEE Intelligent Systems 13:44-49, 1998.

[Yang2016] Z. Yang, T. Zhang, D. Zhang, A novel algorithm with differential evolution and coral reef optimization for extreme learning machine training, Cognitive Neurodynamics 10(1):73-83, 2016.

[Yawei2016] Q. Yawei, T. Na, J. Zhicheng, W. Yan, Coral Reefs Optimization for solving parameter identification in permanent magnet synchronous motor. Journal of System Simulation 28(4): 2016.

[Yiou2014] P. Yiou, M. Boichu, R. Vautard, M. Vrac, S. Jourdain, E. Garnier, F. Fluteau, L. Menut, Ensemble meteorological reconstruction using circulation analogues of 1781-1785. Climate of the Past 10:797-809, 2014.

[Zengand2013] J. Zengand, W. Qiao, Short-term solar power prediction using a support vector machine. Renewable Energy 52:118-127, 2013.

[Zhang2013] W. Zhang, J. Wang, J. Wang, Z. Zhao, M. Tian, Short-term wind speed forecasting based on a hybrid model, Applied Soft Computing 13(7):3225-3233, 2013.

[Zhang2016] C. Zhang, H. Wei, J. Zhao, T. Liu, T. Zhu, K. Zhang, Short-term wind speed forecasting using empirical mode decomposition and feature selection. Renewable Energy 96:727-737, 2016.

[Zhang2017] C. Zhang, J. Zhou, C. Li, W. Fu, T. Peng, A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting. Energy Conversion and Management 143:360-376, 2017.

[Zheng2017] W. Zheng, X. Peng, D. Lu, D. Zhang, Y. Liu, Z. Lin, L. Lin, Composite quantile regression extreme learning machine with feature selection for short-term wind speed forecasting: A new approach. Energy Conversion and Management 151:737-752, 2017.